

IJCSIS Vol. 14 No. 3, March 2016 Part I
ISSN 1947-5500

**International Journal of
Computer Science
& Information Security**

© IJCSIS PUBLICATION 2016
Pennsylvania, USA

Indexed and technically co-sponsored by :



AUTHOR SERIES



IJCSIS

ISSN (online): 1947-5500

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

CALL FOR PAPERS

International Journal of Computer Science and Information Security (IJCSIS) January-December 2016 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.

See authors guide for manuscript preparation and submission guidelines.

Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Scopus Database, Cornell University Library, ScientificCommons, ProQuest, EBSCO and more.

Deadline: see web site

Notification: see web site

Revision: see web site

Publication: see web site

Context-aware systems
Networking technologies
Security in network, systems, and applications
Evolutionary computation
Industrial systems
Evolutionary computation
Autonomic and autonomous systems
Bio-technologies
Knowledge data systems
Mobile and distance education
Intelligent techniques, logics and systems
Knowledge processing
Information technologies
Internet and web technologies
Digital information processing
Cognitive science and knowledge

Agent-based systems
Mobility and multimedia systems
Systems performance
Networking and telecommunications
Software development and deployment
Knowledge virtualization
Systems and networks on the chip
Knowledge for global defense
Information Systems [IS]
IPv6 Today - Technology and deployment
Modeling
Software Engineering
Optimization
Complexity
Natural Language Processing
Speech Synthesis
Data Mining

For more topics, please see web site <https://sites.google.com/site/ijcsis/>

arXiv.org

Google scholar

SCIRUS
search engine for science

ScientificCommons

Scribd

.docstoc
find and share professional documents

BASE
Bielefeld Academic Search Engine

CiteSeer^x beta

dblp.uni-trier.de
Computer Science
Bibliography

DOAJ
DIRECTORY OF
OPEN ACCESS
JOURNALS



ProQuest

For more information, please visit the journal website (<https://sites.google.com/site/ijcsis/>)

Editorial

Message from Editorial Board

It is our great pleasure to present the **March 2016 issue** (Volume 14 Number 3) of the **International Journal of Computer Science and Information Security (IJCSIS)**. High quality survey and review articles are proposed from experts in the field, promoting insight and understanding of the state of the art, and trends in computer science and technology. The contents include original research and innovative applications from all parts of the world. According to Google Scholar, up to now papers published in IJCSIS have been cited over 5668 times and the number is quickly increasing. This statistics shows that IJCSIS has established the first step to be an international and prestigious journal in the field of Computer Science and Information Security. The main objective is to disseminate new knowledge and latest research for the benefit of all, ranging from academia and professional communities to industry professionals. It especially provides a platform for high-caliber researchers, practitioners and PhD/Doctoral graduates to publish completed work and latest development in active research areas. IJCSIS is indexed in major academic/scientific databases and repositories: Google Scholar, CiteSeerX, Cornell's University Library, Ei Compendex, ISI Scopus, DBLP, DOAJ, ProQuest, Thomson Reuters, ArXiv, ResearchGate, Academia.edu and EBSCO among others.

On behalf of IJCSIS community and the sponsors, we congratulate the authors and thank the reviewers for their dedicated services to review and recommend high quality papers for publication. In particular, we would like to thank the international academia and researchers for continued support by citing papers published in IJCSIS. Without their sustained and unselfish commitments, IJCSIS would not have achieved its current premier status.

"We support researchers to succeed by providing high visibility & impact value, prestige and excellence in research publication." For further questions or other suggestions please do not hesitate to contact us at ijcsiseditor@gmail.com.

A complete list of journals can be found at:
<http://sites.google.com/site/ijcsis/>

IJCSIS Vol. 14, No. 3, March 2016 Edition

ISSN 1947-5500 © IJCSIS, USA.

Journal Indexed by (among others):



Open Access This Journal is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source.



Bibliographic Information

ISSN: 1947-5500

Monthly publication (Regular Special Issues)
Commenced Publication since May 2009

Editorial / Paper Submissions:

IJCSIS Managing Editor

ijcsiseditor@gmail.com

Pennsylvania, USA

Tel: +1 412 390 5159

IJCSIS EDITORIAL BOARD

Editorial Board Members	Guest Editors / Associate Editors
Dr. Shimon K. Modi [Profile] Director of Research BSPA Labs, Purdue University, USA	Dr Riktesh Srivastava [Profile] Associate Professor, Information Systems, Skyline University College, Sharjah, PO 1797, UAE
Professor Ying Yang , PhD. [Profile] Computer Science Department, Yale University, USA	Dr. Jianguo Ding [Profile] Norwegian University of Science and Technology (NTNU), Norway
Professor Hamid Reza Naji , PhD. [Profile] Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran	Dr. Naseer Alquraishi [Profile] University of Wasit, Iraq
Professor Yong Li , PhD. [Profile] School of Electronic and Information Engineering, Beijing Jiaotong University, P. R. China	Dr. Kai Cong [Profile] Intel Corporation, & Computer Science Department, Portland State University, USA
Professor Mokhtar Beldjehem , PhD. [Profile] Sainte-Anne University, Halifax, NS, Canada	Dr. Omar A. Alzubi [Profile] Prince Abdullah Bin Ghazi Faculty of Information Technology Al-Balqa Applied University (BAU), Jordan
Professor Yousef Farhaoui , PhD. Department of Computer Science, Moulay Ismail University, Morocco	Dr. Jorge A. Ruiz-Vanoye [Profile] Universidad Autónoma del Estado de Morelos, Mexico
Dr. Alex Pappachen James [Profile] Queensland Micro-nanotechnology center, Griffith University, Australia	Prof. Ning Xu , Wuhan University of Technology, China
Professor Sanjay Jasola [Profile] Dean, School of Information and Communication Technology, Gautam Buddha University	Dr. Bilal Alatas [Profile] Department of Software Engineering, Firat University, Turkey
Dr. Siddhivinayak Kulkarni [Profile] University of Ballarat, Ballarat, Victoria, Australia	Dr. Ioannis V. Koskosas , University of Western Macedonia, Greece
Dr. Reza Ebrahimi Atani [Profile] University of Guilan, Iran	Dr Venu Kuthadi [Profile] University of Johannesburg, Johannesburg, RSA
Dr. Umar Ruhi [Profile] University of Ottawa, Canada	Dr. Zhihan Iv [Profile] Chinese Academy of Science, China
Dr. Vahid Esmaeelzadeh [Profile] Iran University of Science and Technology	Prof. Ghulam Qasim [Profile] University of Engineering and Technology, Peshawar, Pakistan
Dr. Jiliang Zhang [Profile] Northeastern University, China	Prof. Dr. Maqbool Uddin Shaikh [Profile] Preston University, Islamabad, Pakistan
Dr. Jacek M. Czerniak [Profile] Casimir the Great University in Bydgoszcz, Poland	Dr. Musa PEKER [Profile] Faculty of Technology, Mugla Sitki Kocman University, Turkey
	Dr. Wencan Luo [Profile] University of Pittsburgh, US

TABLE OF CONTENTS

1. Paper 290216998: PSNR and Jitter Analysis of Routing Protocols for Video Streaming in Sparse MANET Networks, using NS2 and the Evalvid Framework (pp. 1-9)

Sabrina Nefti, Dept. of Computer Science, University Batna 2, Algeria
Mammar Sedrati, Dept. of Computer Science, University Batna 2, Algeria

Abstract — Advances in multimedia and ad-hoc networking have urged a wealth of research in multimedia delivery over ad-hoc networks. This comes as no surprise, as those networks are versatile and beneficial to a plethora of applications where the use of fully wired network has proved intricate if not impossible, such as prompt formation of networks during conferences, disaster relief in case of flood and earthquake, and also in war activities. In this paper, we aim to investigate the combined impact of network sparsity and network node density on the Peak Signal Noise to Ratio (PSNR) and jitter performance of proactive and reactive routing protocols in ad-hoc networks. We also shed light onto the combined effect of mobility and sparsity on the performance of these protocols. We validate our results through the use of an integrated Simulator-Evaluator environment consisting of the Network Simulator NS2, and the Video Evaluation Framework Evalvid.

Keywords- PSNR, MANET, Sparsity, Density, Routing protocols, Video Streaming, NS2, Evalvid

2. Paper 290216996: Automatically Determining the Location and Length of Coronary Artery Thrombosis Using Coronary Angiography (pp. 10-19)

Mahmoud Al-Ayyoub, Ala'a Oqaily and Mohammad I. Jarrah
Jordan University of Science and Technology Irbid, Jordan
Huda Karajeh, The University of Jordan Amman, Jordan

Abstract — Computer-aided diagnosis (CAD) systems have gained a lot of popularity in the past few decades due to their effectiveness and usefulness. A large number of such systems are proposed for a wide variety of abnormalities including those related to coronary artery disease. In this work, a CAD system is proposed for such a purpose. Specifically, the proposed system determines the location of thrombosis in x-ray coronary angiograms. The problem at hand is a challenging one as indicated by some researchers. In fact, no prior work has attempted to address this problem to the best of our knowledge. The proposed system consists of four stages: image preprocessing (which involves noise removal), vessel enhancement, segmentation (which is followed by morphological operations) and localization of thrombosis (which involves skeletonization and pruning before localization). The proposed system is tested on a rather small dataset and the results are encouraging with a 90% accuracy.

Keywords — Heterogeneous wireless networks, Vertical handoff, Markov model, Artificial intelligence, Mobility management.

3. Paper 29021671: Neutralizing Vulnerabilities in Android: A Process and an Experience Report (pp. 20-29)

Carlos André Batista de Carvalho (#), Rossana Maria de Castro Andrade (*), Márcio E. F. Maia (*), Davi Medeiros Albuquerque (*), Edgar Tarton Oliveira Pedrosa (*)*
Computer Science Department, Federal University of Piauí, Brazil
** Group of Computer Networks, Software Engineering, and Systems, Federal University of Ceará, Brazil*

Abstract — Mobile devices became a natural target of security threats due their vast popularization. That problem is even more severe when considering Android platform, the market leader operating system, built to be open and extensible. Although Android provides security countermeasures to handle mobile threats, these defense measures are not sufficient and attacks can be performed in this platform, exploiting existing vulnerabilities. Then, this paper

focuses on improving the security of the Android ecosystem with a contribution that is two-fold, as follows: i) a process to analyze and mitigate Android vulnerabilities, scrutinizing existing security breaches found in the literature and proposing mitigation actions to fix them; and ii) an experience report that describes four vulnerabilities and their corrections, being one of them a new detected and mitigated vulnerability.

4. Paper 29021655: Performance Analysis of Proposed Network Architecture: OpenFlow vs. Traditional Network (pp. 30-39)

Idris Z. Bholebawa (#), Rakesh Kumar Jha (), Upena D. Dalal (#)*

(#) Department of Electronics and Communication Engineering, S. V. National Institute of Technology, Surat, Gujarat, India.

() School of Electronics and Communication Engineering, Shri Mata Vaishno Devi University, Katra, J&K*

Abstract – The Internet has been grown up rapidly and supports variety of applications on basis of user demands. Due to emerging technological trends in networking, more users are becoming part of a digital society, this will ultimately increases their demands in diverse ways. Moreover, traditional IP-based networks are complex and somehow difficult to manage because of vertical integration problem of network core devices. Many research projects are under deployment in this particular area by network engineers to overcome difficulties of traditional network architecture and to fulfill user requirements efficiently. A recent and most popular network architecture proposed is Software-Defined Networks (SDN). A purpose of SDN is to control data flows centrally by decoupling control plane and data plane from network core devices. This will eliminate the difficulty of vertical integration in traditional networks and makes the network programmable. A most successful deployment of SDN is OpenFlow-enabled networks.

In this paper, a comparative performance analysis between traditional network and OpenFlow-enabled network is done. A performance analysis for basic and proposed network topologies is done by comparing round-trip propagation delay between end nodes and maximum obtained throughput between nodes in traditional and OpenFlow-enabled network environment. A small campus network have been proposed and performance comparison between traditional network and OpenFlow-enabled network is done in later part of this paper. An OpenFlow-enabled campus network is proposed by interfacing virtual node of virtually created OpenFlow network with real nodes available in campus network. An implementation of all the OpenFlow-enabled network topologies and a proposed OpenFlow-enabled campus network is done using open source network simulator and emulator called Mininet. All the traditional network topologies are designed and analyzed using NS2 - network simulator.

Keywords – *SDN, OpenFlow, Mininet, Network Topologies, Interfacing Network.*

5. Paper 29021622: Reverse Program Analyzed with UML Starting from Object Oriented Relationships (pp. 40-45)

Hamed J. Al-Fawareh, Software Engineering Department, Zarka University, Jordan

Abstract - In this paper, we provide a reverse-tool for object oriented programs. The tool focuses on the technical side of maintaining object-oriented program and the description of associations graph for representing meaningful diagram between components of object-oriented programs. In software maintenance perspective reverse engineering process extracts information to provide visibility of the object oriented components and relations in the software that are essential for maintainers.

Keywords: *Software Maintenance, Reverse Engineering.*

6. Paper 29021628: Lifetime Optimization in Wireless Sensor Networks Using FDstar-Lite Routing Algorithm (pp. 46-55)

Imad S. Alshawi, College of Computer Science and Information Technology, Basra University, Basra, Iraq

Ismaiel O. Alalewi, College of Science, Basra University, Basra, Iraq

Abstract — Commonly in Wireless Sensor Networks (WSNs), the biggest challenge is to make sensor nodes that are energized by low-cost batteries with limited power run for longest possible time. Thus, energy saving is indispensable concept in WSNs. The method of data routing has a pivotal role in conserving the available energy since remarkable amount of energy is consumed by wireless data transmission. Therefore, energy efficient routing protocols can save battery power and give the network longer lifetime. Using complex protocols to plan data routing efficiently can reduce energy consumption but can produce processing delay. This paper proposes a new routing method called FDstar-Lite which combines Dstar-Lite algorithm with Fuzzy Logic. It is used to find the optimal path from the source node to the destination (sink) and reuse that path in such a way that keeps energy consumption fairly distributed over the nodes of a WSN while reducing the delay of finding the routing path from scratch each time. Interestingly, FDstar-Lite was observed to be more efficient in terms of reducing energy consumption and decreasing end-to-end delay when compared with A-star algorithm, Fuzzy Logic, Dstar-Lite algorithm and Fuzzy A-star. The results also show that, the network lifetime achieved by FDstar-Lite could be increased by nearly 35%, 31%, 13% and 11% more than that obtained by A-star algorithm, Fuzzy Logic, Dstar-Lite algorithm and Fuzzy A-star respectively.

Keywords— *Dstar-Lite algorithm, fuzzy logic, network lifetime, routing, wireless sensor network.*

7. Paper 29021637: An Algorithm for Signature Recognition Based on Image Processing and Neural Networks (pp. 56-60)

Ramin Dehgani, Ali Habiboghli

Department of computer science and engineering, Islamic Azad University, Khoy, Iran

Abstract — Characteristics related to people signature has been extracted in this paper. Extracted Specialty vector under neural network has been used for education. After teaching network, signatures have been evaluated by educated network to recognize real signature from unreal one. Comparing the results shows that the efficiency of this method is better than the other methods.

Index Terms— *signature recognition, neural networks, image processing.*

8. Paper 29021640: A Report on Using GIS in Establishing Electronic Government in Iraq (pp. 61-64)

Ahmed M.JAMEL, Department of Computer Engineering, Erciyes University, Kayseri, Turkey

Dr. Tolga PUSATLI, Department of Mathematics and Computer Science, Cankaya University, Ankara, Turkey

Abstract — Electronic government initiatives and public participation in them are among the indicators of today's development criteria for countries. After the consequent of two wars, Iraq's current position in, for example, the UN's e-government ranking is quite low and did not improve in recent years. In the preparation of this work, we are motivated by the fact that handling geographic data of the public facilities and resources are needed in most of the e-government projects. Geographical information systems (GIS) provide the most common tools, not only to manage spatial data, but also to integrate with non-spatial attributes of the features. This paper proposes that establishing a working GIS in the health sector of Iraq would improve e-government applications. As a case study, investigating hospital locations in Erbil has been chosen. It is concluded that not much is needed to start building base works for GIS supported e-government initiatives.

Keywords - *Electronic government, Iraq, Erbil, GIS, Health Sector.*

9. Paper 29021642: Satellite Image Classification by Using Distance Metric (pp. 65-68)

Dr. Salem Saleh Ahmed Alamri, Dr. Ali Salem Ali Bin-Sama

Department of Engineering Geology, Oil & Minerals Faculty, Aden University, Aden, Yemen

Dr. Abdulaziz Saleh Yeslam Bin-Habtoor

Department of Electronic and Communication Engineering, Faculty of Engineering & Petroleum, Hadramote University, Mokula, Yemen

Abstract — This paper attempts to undertake the study satellite image classification by using six distance metric as Bray Curtis Distance Method, Canberra Distance Method, Euclidean Distance Method, Manhattan Distance Method, Square Chi Distance Method, Squared Chord Distance Method and they are compared with one another, So as to choose the best method for satellite image classification.

Keyword: Satellite Image, Classification, Texture Image, Distance Metric,

10. Paper 29021650: Cybercrime and its Impact on E-government Services and the Private Sector in The Middle East (pp. 69-73)

Sulaiman Al Amro, Computer Science (CS) Department, Qassim University, Buraydah, Qassim, 51452, KSA

Abstract — This paper will discuss the issue of cybercrime and its impact on both e-government services and the private sector in the Middle East. The population of the Middle East has now become increasingly connected, with ever greater use of technology. However, the issue of piracy has continued to escalate, without any signs of abating. Acts of piracy have been established as the most rapidly growing (and efficient) sector within the Middle East, taking advantage of attacks on the infrastructure of information technology. The production of malicious software and new methods of breaching security has enabled both amateur and professional hackers and spammers, etc., to target the Internet in new and innovative ways, which are, in many respects, similar to legitimate businesses in the region.

Keywords - cybercrimes; government sector; private sectors; Middle East; computer security

11. Paper 29021657: Performance Comparison between Forward and Backward Chaining Rule Based Expert System Approaches Over Global Stock Exchanges (pp. 74-81)

Sachin Kamley, Deptt. of Computer Application's S.A.T.I., Vidisha, India

Shailesh Jaloree, Deptt. of Appl. Math's and CS S.A.T.I., Vidisha, India

R.S. Thakur, Deptt. of Computer Application's M.A.N.I.T., Bhopal, India

Abstract — For the last couple of decade's stock market has been considered as a most noticeable research area everywhere throughout the world because of the quickly developing of the economy. Throughout the years, a large portion of the researchers and business analysts have been contributed around there. Extraordinarily, Artificial Intelligence (AI) is the principle overwhelming area of this field. In AI, an expert system is one of the understood and prevalent techniques that copy the human abilities in order to take care of particular issues. In this research study, forward and backward chaining two primary expert system inference methodologies is proposed to stock market issue and Common LISP 3.0 based editors are used for designing an expert system shell. Furthermore, expert systems are tested on four noteworthy global stock exchanges, for example, India, China, Japan and United States (US). In addition, different financial components, for example, Gross Domestic Product (GDP), Unemployment Rate, Inflation Rate and Interest Rate are also considered to build the expert knowledge base system. Finally, experimental results demonstrate that the backward chaining approach has preferable execution performance over forward chaining approach.

Keywords— Stock Market; Artificial Intelligence; Expert System; Macroeconomic Factors; Forward Chaining; Backward Chaining; Common LISP 3.0.

12. Paper 29021658: Analysis of Impact of Varying CBR Traffic with OLSR & ZRP (pp. 82-85)

Rakhi Purohit, Department of Computer Science & Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India

Bright Keswani, Associate Professor & Head Department of Computer Application, Suresh Gyan Vihar University, Jaipur, Rajasthan, India

Abstract — Mobile ad hoc network is the way to interconnect various independent nodes. This network is decentralize and not follows any fixed infrastructure. All the routing functionality are controlled by all the nodes. Here nodes can be volatile in nature so they can change place in network and effect network architecture. Routing in mobile ad hoc network is very much dependent on its protocols which can be proactive and reactive as well as with both features. This work consist of analysis of protocols have analyzed in different scenarios with varying data traffic in the network. Here OLSR protocol has taken as proactive and ZRP as Hybrid protocol. Some of the calculation metrics have evaluated for this analysis. This analysis has performed on well-known network simulator NS2.

Index Terms:- Mobile ad hoc network, Routing, OLSR, Simulation, and NS2.

13. Paper 29021662: Current Moroccan Trends in Social Networks (pp. 86-98)

Abdeljalil EL ABDOULI, Abdelmajid CHAFFAI, Larbi HASSOUNI, Houda ANOUN, Khalid RIFI, RITM Laboratory, CED Engineering Sciences, Ecole Supérieure de Technologie, Hassan II University of Casablanca, Morocco

Abstract — The rapid development of social networks during the past decade has lead to the emergence of new forms of communication and new platforms like Twitter and Facebook. These are the two most popular social networks in Morocco. Therefore, analyzing these platforms can help in the interpretation of Moroccan society current trends. However, this will come with few challenges. First, Moroccans use multiple languages and dialects for their daily communication, such as Standard Arabic, Moroccan Arabic called “Darija”, Moroccan Amazigh dialect called “Tamazight”, French, and English. Second, Moroccans use reduced syntactic structures, and unorthodox lexical forms, with many abbreviations, URLs, #hashtags, spelling mistakes. In this paper, we propose a detection engine of Moroccan social trends, which can extract the data automatically, store it in a distributed system which is the Framework Hadoop using the HDFS storage model. Then we process this data, and analyze it by writing a distributed program with Pig UDF using Python language, based on Natural Language Processing (NLP) as linguistic technique, and by applying the Latent Dirichlet Allocation (LDA) for topic modeling. Finally, our results are visualized using pyLDAvis, WordCloud, and exploratory data analysis is done using hierarchical clustering and other analysis methods.

Keywords: distributed system; Framework Hadoop; Pig UDF; Natural Language Processing; Latent Dirichlet Allocation; topic modeling; pyLDAvis; wordcloud; exploratory data analysis; hierarchical clustering.

14. Paper 29021665: Design Pattern for Multilingual Web System Development (pp. 99-105)

Dr. Habes Alkhraisat, Al Balqa Applied University, Jordan

Abstract — Recently- Multilingual WEB Database system have brought into sharp focus the need for systems to store and manipulate text data efficiently in a suite of natural languages. While some means of storing and querying multilingual data are provided by all current database systems. In this paper, we present an approach for efficient development multilingual web database system with the use of object oriented design principle benefits. We propose functional, efficient, dynamic and flexible object oriented design pattern and database system architecture for making the performance of the database system to be language independent. Results from our initial implementation of the proposed methodology are encouraging indicating the value of proposed approach.

Index Terms— Database System, Design Pattern, Inheritance, Object Oriented, Structured Query Language.

15. Paper 29021669: A Model for Deriving Matching Threshold in Fingerprint-based Identity Verification System (pp. 106-114)

Omolade Ariyo. O., Fatai Olawale. W. Department of Computer Science, University of Ilorin, Ilorin, Nigeria

Abstract - Currently there is a variety of designs and Implementation of biometric especially fingerprint. There is currently a standard used for determining matching threshold, which allows vendors to skew their test results in their favour by using assumed figure between -1 to +1 or values between 1 and 100%. The research contribution in this research work is to formulate an equation to determine the threshold against which the minutia matching score will be compare using the features set of the finger itself which is devoid of assumptions. Based on the results of this research, it shows that the proposed design and development of a fingerprint-based identity verification system can be achieved without riding on assumptions. Thereby, eliminating the false rate of Acceptance and reduce false rate of rejection as a result of the threshold computation using the features of the enrolled finger. Further research can be carried out in the area of comparing matching result generated from the threshold assumption with the threshold computation formulated in this thesis paper.

Keywords: Biometrics; Threshold; Matching; Algorithm; Scoring.

16. Paper 29021682: A Sliding Mode Controller for Urea Plant (pp. 115-126)

*M. M. Saafan, M. M. Abdelsalam, M. S. Elksasy, S. F. Saraya, and F. F.G. Areed
Computers and Control Systems Engineering Department, Faculty of Engineering, Mansoura University, Egypt.*

Abstract - The present paper introduces the mathematical model of urea plant and suggests two methods for designing special purpose controllers. The first proposed method is PID controller and the second is sliding mode controller (SMC). These controllers are applied for a multivariable nonlinear system as a Urea Reactor system. The main target of the designed controllers is to reduce the disturbance of NH₃ pump and CO₂ compressor in order to reduce the pollution effect in such chemical plant. Simulation results of the suggested PID controller are compared with that of the SMC controller. Comparative analysis proves the effectiveness of the suggested SMC controller than the PID controller according to disturbance minimization as well as dynamic response. Also, the paper presents the results of applying SMC, while maximizing the production of the urea by maximizing the NH₃ flow rate. This controller kept the reactor temperature, the reactor pressure, and NH₃/CO₂ ratio in the suitable operating range. Moreover, the suggested SMC when compared with other controllers in the literature shows great success in maximizing the production of urea.

Keywords: Sliding mode controller, PID controller, urea reactor, Process Control, Chemical Industry, Adaptive controller, Nonlinearity.

17. Paper 29021683: Transmission Control Protocol and Congestion Control: A Review of TCP Variants (pp. 127-135)

*Babatunde O. Olasoji, Oyenike Mary Olanrewaju, Isaiah O. Adebayo
Mathematical Sciences and Information Technology Department, Federal University Dutsinma, Katsina State, Nigeria.*

Abstract - Transmission control protocol (TCP) provides a reliable data transfer in all end-to-end data stream services on the internet. There are some mechanisms that TCP has that make it suitable for this purpose. Over the years, there have been modifications in TCP algorithms starting from the basic TCP that has only slow-start and congestion avoidance algorithm to the modifications and additions of new algorithms. Today, TCP comes in various variants which include TCP Tahoe, Reno, new Reno, Vegas, sack etc. Each of this TCP variant has its peculiarities, merits and demerits. This paper is a review of four TCP variants, they are: TCP Tahoe, Reno, new Reno and Vegas, their congestion avoidance algorithms, and possible future research areas.

Keywords – Transmission control protocol; Congestion Control; TCP Tahoe; TCP Reno; TCP New Reno; TCP Vegas

18. Paper 31011656: Detection of Black Hole Attacks in MANETs by Using Proximity Set Method (pp. 136-145)

K. Vijaya Kumar, Research Scholar (Karpagam University), Assistant Professor, Department of Computer Science Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India.

Dr. K. Somasundaram, Professor, Department of Computer Science and Engg., Vel Tech High Tech Dr.RR Dr.SR Engineering College, Avadi, Chennai, Tamilnadu India

Abstract - A Mobile Adhoc Networks (MANETS) is an infrastructure less or self-configuring network which contain a collection of mobile nodes moving randomly by changing their topology with limited resources. These Networks are prone to different types of attacks due to lack of central monitoring facility. The main aim is to inspect the effect of black hole attack on the network layer of MANET. A black hole attack is a network layer attack also called sequence number attack which utilizes the destination sequence number to claim that it has a shortest route to reach the destination and consumes all the packets forwarded by the source. To diminish the effects of such attack, we have proposed a detection technique by using Proximity Set Method (PSM) that efficiently detects the malicious nodes in the network. The severity of attack depends on the position of the malicious node that is near, midway or far from the source. The various network scenarios of MANETS with AODV routing protocol are simulated using NS2 simulator to analyze the performance with and without the black hole attack. The performance parameters like PDR, delay, throughput, packet drop and energy consumption are measured. The overall throughput and PDR increases with the number of flows but reduces with the attack. With the increase in the black hole attackers, the PDR and throughput reduces and close to zero as the number of black hole nodes are maximum. The packet drop also increases with the attack. The overall delay factor varies based on the position of the attackers. As the mobility varies the delay and packet drop increases but PDR and throughput decreases as the nodes moves randomly in all directions. Finally the simulation results gives a very good comparison of performance of MANETS with original AODV, with black hole attack and applying proximity set method for presence of black hole nodes different network scenarios.

Keywords: AODV protocol, security, black hole attack, NS2 simulator, proximity set method, performance parameters.

19. Paper 290216995: A Greedy Approach to Out-Door WLAN Coverage Planning (pp. 146-152)

Gilbert M. Gilbert, College of Informatics and Virtual Education, The University of Dodoma

Abstract — Planning for optimal out-door wireless network coverage is one of the core issues in network design. This paper considers coverage problem in outdoor-wireless networks design with the main objective of proposing methods that offer near-optimal coverage. The study makes use of the greedy algorithms and some specified criteria (field strength) to find minimum number of base stations and access points that can be activated to provide maximum services (coverage) to a specified number of users. Various wireless network coverage planning scenarios were considered to an imaginary town subdivided into areas and a comprehensive comparison among them was done to offer desired network coverage that meet the objective.

Keywords — greedy algorithms, outdoor-wlan, coverage planning, greedy algorithms, path loss.

20. Paper 29021620: Cerebellar Model Articulation Controller Network for Segmentation of Computer Tomography Lung Image (pp. 153-157)

(1) Benita K.J. Veronica, (2) Purushothaman S., Rajeswari P.,

(1) Mother Teresa Women's University, Kodaikanal, India.

(2) Associate Professor, Institute of Technology, Haramaya University, Ethiopia.

Abstract - This paper presents the implementation of CMAC network for segmentation of computed tomography lung slice. Representative features are extracted from the slice to train the CMAC algorithm. At the end of training, the final weights are stored in the database. During the testing the CMAC, a lung slice is presented to obtain the segmented image.

Keywords: CMAC; segmentation; computed tomography; lung slice

21. Paper 29021625: Performance Evaluation of Pilot-Aided Channel Estimation for MIMO-OFDM Systems (pp. 158-162)

*B. Soma Sekhar, Dept of ECE, Sanketika Vidhya Parishad Engg. College, Visakhapatnam, Andhra Pradesh, India
A. Mallikarjuna Prasad, Dept of ECE, University College of Engineering, Kakinada, JNTUK, Andhra Pradesh, India*

Abstract — In this paper a pilot aided channel estimation for Multiple-Input Multiple-Output/Orthogonal Frequency-Division Multiplexing (MIMO/ OFDM) systems in time-varying wireless channels is considered. Channel coefficients can be modeled by using truncated discrete Fourier Basis Expansion model (Fourier-BEM) and a discrete prolate spheroidal sequence model (DPSS). The channel is assumed which is varying linearly with respect to time. Based on these models, a weighted average approach is adopted for estimating LTV channels for OFDM symbols. The performance analysis between Fourier BEM, DPSS models, Legendre and Chebishev polynomial based on Mean square error (MSE) is present. Simulation results show that the DPSS-BEM model outperforms the Fourier Basis expansion model.

Index Terms — *Basis Expansion Model (BEM), Discrete Prolate Spheroidal Sequence (DPSS), Mean Square Error (MSE).*

22. Paper 29021674: Investigating the Distributed Load Balancing Approach for OTIS-Star Topology (pp. 163-171)

Ahmad M. Awwad, Jehad Al-Sadi

Abstract — This research effort investigates and proposes an efficient method for load balancing problem for the OTIS-Star topology. The proposed method is named OTIS-Star Electronic-Optical-Electronic Exchange Method; OSEOEM; which utilizes the electronic and optical technologies facilitated by the OTIS-Star topology. This method is based on the previous FOFEM algorithm for OTIS-Cube networks. A complete investigation of the OSEOEM is introduced in this paper including a description of the algorithm and the stages of performing Load Balancing. A comprehensive analytical and theoretical study to prove the efficiency of this method, and statistical outcomes based on common used performance measures has been also presented. The outcome of this investigation proves the efficiency of the proposed OSEOEM method.

Keywords — *Electronic Interconnection Networks, Optical Networks, Load balancing, Parallel Algorithms, OTIS-Star Network.*

23. Paper 29021694: American Sign Language Pattern Recognition Based on Dynamic Bayesian Network (pp. 172-177)

*Habes Alkhraisat, Saqer Alshrah
Department of Computer Science, Al-Balqa Applied University, Jordan*

Abstract — Sign languages are usually developed among deaf communities, which include friends and families of deaf people or people with hearing impairment. American Sign Language (ASL) is the primary language used by the American Deaf Community. It is not simply a signed representation of English, but rather, a rich natural language with a unique structure, vocabulary, and grammar. In this paper, we propose a method for American Sign Language alphabet, and number gestures interpretation in a continuous video stream using a dynamic Bayesian network. The experimental result, using RWTHBOSTON-104 data set, shows a recognition rate upwards of 99.09%.

Index Terms — *American Sign Language (ASL), Dynamic Bayesian Network, Hand Tracking, Feature extraction.*

24. Paper 2902169910: Identification of Breast Cancer by Artificial Bee Colony Algorithm with Least Square Support Vector Machine (pp. 178-183)

S. Mythili, PG & Research Department of Computer Application, Hindusthan College of Arts & Science, Coimbatore, India

Dr. A. V. Senthilkumar, Director, PG & Research Department of Computer Application, Hindusthan College of Arts & Science, Coimbatore, India

Abstract - Procedure for the identification of several discriminant factors. A new method is proposed for identification of Breast Cancer in Peripheral Blood with microarray Datasets by introducing the Hybrid Artificial Bee Colony (ABC) algorithm with Least Squares Support Vector Machine (LS-SVM), namely as ABC-SVM. Breast cancer is identified by Circulating Tumor Cells in the Peripheral Blood. The mechanisms that implicate Circulating Tumor Cells (CTC) in metastatic disease is notably in Metastatic Breast Cancer (MBC), remain elusive. The proposed work is focused on the identification of tissues in Peripheral Blood that can indirectly reveal the presence of cancer cells. By selecting publicly available Breast Cancer tissues and Peripheral Blood microarray datasets, we follow two-step elimination.

Keywords: Breast Cancer (BC), Circulating Tumor Cells (CTC), Peripheral Blood (PB), Artificial Bee Colony (ABC), Least Squares Support Vector Machine (LSSVM).

25. Paper 29021601: Moving Object Segmentation and Vibrant Background Elimination Using LS-SVM (pp. 184-197)

Mehul C. Parikh, Computer Engineering Department, Charotar University of Science and Technology, Changa, Gujarat, India.

Kishor G. Maradia, Department of Electronics and Communication, Government Engineering College, Gandhinagar, Gujarat, India

Abstract - Moving object segmentation is a significant research area in the field of computer intelligence due to technological and theoretical progress. Many approaches are being developed for moving object segmentation. These approaches are useful for specific situation but have many restrictions. Execution speed of these approaches is one of the major limitations. Machine learning techniques are used to decrease time and improve quality of result. LS-SVM optimizes result quality and time complexity in classification problem. This paper describes an approach to segment moving object and vibrant background elimination using the least squares support vector machine method. In this method consecutive frame difference was given as an input to bank of Gabor filter to detect texture feature using pixel intensity. Mean value of intensity on $4 * 4$ block of image and on whole image was calculated and which are then used to train LS-SVM model using random sampling. Trained LS-SVM model was then used to segment moving object from the image other than the training images. Results obtained by this approach are very promising with improvement in execution time.

Key Words: Segmentation, Machine Learning, Gabor filter, LS-SVM.

26. Paper 29021609: On Annotation of Video Content for Multimedia Retrieval and Sharing (pp. 198-218)

Mumtaz Khan, Shah Khusro, Irfan Ullah

Department of Computer Science, University of Peshawar, Peshawar 25120, Pakistan

Abstract - The development of standards like MPEG-7, MPEG-21 and ID3 tags in MP3 have been recognized from quite some time. It is of great importance in adding descriptions to multimedia content for better organization and retrieval. However, these standards are only suitable for closed-world-multimedia-content where a lot of effort is put in the production stage. Video content on the Web, on the contrary, is of arbitrary nature captured and uploaded in a variety of formats with main aim of sharing quickly and with ease. The advent of Web 2.0 has resulted in the wide availability of different video-sharing applications such as YouTube which have made video as major content on the

Web. These web applications not only allow users to browse and search multimedia content but also add comments and annotations that provide an opportunity to store the miscellaneous information and thought-provoking statements from users all over the world. However, these annotations have not been exploited to their fullest for the purpose of searching and retrieval. Video indexing, retrieval, ranking and recommendations will become more efficient by making these annotations machine-processable. Moreover, associating annotations with a specific region or temporal duration of a video will result in fast retrieval of required video scene. This paper investigates state-of-the-art desktop and Web-based-multimedia annotation-systems focusing on their distinct characteristics, strengths and limitations. Different annotation frameworks, annotation models and multimedia ontologies are also evaluated.

Keywords: Ontology, Annotation, Video sharing web application

27. Paper 29021617: A New Approach for Energy Efficient Linear Cluster Handling Protocol In WSN (pp. 219-227)

Jaspinder Kaur, Varsha Sahni

Abstract - Wireless Sensor Networks (WSN) is a rising field for researchers in the recent years. For obtaining durability of network lifetime, and reducing energy consumption, energy efficiency routing protocol play an important role. In this paper, we present an innovative and energy efficient routing protocol. A New linear cluster handling (LCH) technique towards Energy Efficiency in Linear WSNs with multiple static sinks [4] in a linearly enhanced field of 1500m*350m². We are divided the whole into four equal sub-regions. For efficient data gathering, we place three static sinks i.e. one at the centre and two at the both corners of the field. A reactive and Distance plus energy dependent clustering protocol Threshold Sensitive Energy efficient with Linear Cluster Handling [4] DE (TEEN-LCH) is implemented in the network field. Simulation shows improved results for our proposed protocol as compared to TEEN-LCH, in term of throughput, packet delivery ratio and energy consumption.

Keywords: WSN; Routing Protocol; Throughput; Energy Consumption; Packet Delivery

28. Paper 29021624: Protection against Phishing in Mobile Phones (pp. 228-233)

Avinash Shende, IT Department, SRM University, Kattankulathur, Chennai, India

Prof. D. Saveetha, Assistant Professor, IT Department, SRM University, Kattankulathur, Chennai, India

Abstract - Phishing is the attempt to get confidential information such as user-names, credit card details, passwords and pins, often for malicious reasons, by making people believe that they are communicating with legitimate person or identity. In recent years we have seen increase in threat of phishing on mobile phones. In fact, mobile phone phishing is more dangerous than phishing on desktop because of limitations of mobile phones like mobile user habits and small screen. Existing mechanism made for detecting phishing attacks on computers are not able to avoid phishing attacks on mobile devices. We present an anti-phishing mechanism for mobile devices. Our solution verifies if webpages is legitimate or not by comparing the actual identity of webpage with the claimed identity of the webpage. We will use OCR tool to find the identity claimed by the webpage.

29. Paper 29021626: Hybrid Cryptography Technique for Information Systems (pp. 234-243)

Zohair Malki

Faculty of Computer Science and Engineering, Taibah University, Yanbu, Saudi Arabia

Abstract - Information systems based applications are increasing rapidly in many fields including educational, medical, commercial and military areas, which have posed many security and privacy challenges. The key component of any security solution is encryption. Encryption is used to hide the original message or information in a new form that can be retrieved by the authorized users only. Cryptosystems can be divided into two main types: symmetric and asymmetric systems. In this paper we discussed some common systems that belong to both types. Specifically, we will discuss, compare and test the implementation for RSA, RC5, DES, Blowfish and Twofish. Then, a new hybrid

system composed of RSA and RC5 is proposed and tested against these two systems when each used alone. The obtained results show that the proposed system achieves better performance.

30. Paper 29021629: An Efficient Network Traffic Filtering that Recognize Anomalies with Minimum Error Received (pp. 244-256)

*Mohammed N. Abdul Wahid and Azizol Bin Abdullah
Department of Communication Technology and Networks, Faculty of Computer Science and Information
Technology, University Putra Malaysia, Malaysia*

Abstract - The main method is related to processing and filtering data packets on a network system and, more specifically, analyzing data packets transmitted on a regular speed communications links for errors and attackers' detection and signal integrity analysis. The idea of this research is to use flexible packet filtering which is a combination of both the static and dynamic packet filtering with the margin of support vector machine. Many experiments have been conducted in order to investigate the performance of the proposed schemes and comparing them with recent software's that is most relatively to our proposed method that measuring the bandwidth, time, speed and errors. These experiments are performed and examined under different network environments and circumstances. The comparison has been done and results proved that our method gives less error received from the total analyzed packets.

Keywords: Anomaly Detection, Data Mining, Data Processing, Flexible Packet Filtering, Misuse Detection, Network Traffic Analyzer, Packet sniffer, Support Vector Machine, Traffic Signature Matching, User Profile Filter.

31. Paper 29021634: Proxy Blind Signcryption Based on Elliptic Curve Discrete Logarithm Problem (pp. 257-262)

*Anwar Sadat, Department of Information Technology, kohat University of Science and Technology K-P, Pakistan
Insaf Ullah, Hizbullah Khattak, Sultan Ullah, Amjad-ur-Rehman
Department of Information Technology, Hazara Uuniversity Mansehra K-P, Pakistan*

Abstract - Nowadays anonymity, rights delegations and hiding information play primary role in communications through internet. We proposed a proxy blind signcryption scheme based on elliptic curve discrete logarithm problem (ECDLP) meet all the above requirements. The design scheme is efficient and secure because of elliptic curve crypto system. It meets the security requirements like confidentiality, Message Integrity, Sender public verifiability, Warrant unforgeability, Message Unforgeability, Message Authentication, Proxy Non-Repudiation and blindness. The proposed scheme is best suitable for the devices used in constrained environment.

Keywords: proxy signature, blind signature, elliptic curve, proxy blind signcryption.

32. Paper 29021646: A Comprehensive Survey on Hardware/Software Partitioning Process in Co-Design (pp. 263-279)

*Imene Mhadhbi, Slim BEN OTHMAN, Slim Ben Saoud
Department of Electrical Engineering, National Institute of Applied Sciences and Technology, Polytechnic School of
Tunisia, Advanced Systems Laboratory, B.P. 676, 1080 Tunis Cedex, Tunisia*

Abstract - Co-design methodology deals with the problem of designing complex embedded systems, where Hardware/software partitioning is one key challenge. It decides strategically the system's tasks that will be executed on general purpose units and the ones implemented on dedicated hardware units, based on a set of constraints. Many relevant studies and contributions about the automation techniques of the partitioning step exist. In this work, we explore the concept of the hardware/software partitioning process. We also provide an overview about the historical achievements and highlight the future research directions of this co-design process.

Keywords: Co-design; embedded system; hardware/software partitioning; embedded architecture

33. Paper 29021647: Heterogeneous Embedded Network Evaluation of CAN-Switched ETHERNET Architecture (pp. 280-294)

*Nejla Rejeb, Ahmed Karim Ben Salem, Slim Ben Saoud
LSA Laboratory, INSAT-EPT, University of Carthage, TUNISIA*

Abstract - The modern communication architecture of new generation transportation systems is described as heterogeneous. This new architecture is composed by a high rate Switched ETHERNET backbone and low rate data peripheral buses coupled with switches and gateways. Indeed, Ethernet is perceived as the future network standard for distributed control applications in many different industries: automotive, avionics and industrial automation. It offers higher performance and flexibility over usual control bus systems such as CAN and Flexray. The bridging strategy implemented at the interconnection devices (gateways) presents a key issue in such architecture. The aim of this work consists on the analysis of the previous mixed architecture. This paper presents a simulation of CAN-Switched Ethernet network based on OMNET++. To simulate this network, we have also developed a CAN-Switched Ethernet Gateway simulation model. To analyze the performance of our model we have measured the communication latencies per device and we have focused on the timing impact introduced by various CAN-Ethernet multiplexing strategies at the gateways. The results herein prove that regulating the gateways CAN remote traffic has an impact on the end to end delays of CAN flow. Additionally, we demonstrate that the transmission of CAN data over an Ethernet backbone depends heavily on the way this data is multiplexed into Ethernet frames.

Keywords: Ethernet, CAN, Heterogeneous Embedded networks, Gateway, Simulation, End to end delay.

34. Paper 29021660: Reusability Quality Attributes and Metrics of SaaS from Perspective of Business and Provider (pp. 295-312)

*Areeg Samir, Nagy Ramadan Darwish
Information Technology and System, Institute of Statistical Studies and Research*

Abstract - Software as a Service (SaaS) is defined as a software delivered as a service. SaaS can be seen as a complex solution, aiming at satisfying tenants requirements during runtime. Such requirements can be achieved by providing a modifiable and reusable SaaS to fulfill different needs of tenants. The success of a solution not only depends on how good it achieves the requirements of users but also on modifies and reuses provider's services. Thus, providing reusable SaaS, identifying the effectiveness of reusability and specifying the imprint of customization on the reusability of application still need more enhancements. To tackle these concerns, this paper explores the common SaaS reusability quality attributes and extracts the critical SaaS reusability attributes based on provider side and business value. Moreover, it identifies a set of metrics to each critical quality attribute of SaaS reusability. Critical attributes and their measurements are presented to be a guideline for providers and to emphasize the business side.

Index Terms - *Software as a Service (SaaS), Quality of Service (QoS), Quality attributes, Metrics, Reusability, Customization, Critical attributes, Business, Provider.*

35. Paper 29021661: A Model Driven Regression Testing Pattern for Enhancing Agile Release Management (pp. 313-333)

*Maryam Nooraei Abadeh, Department of Computer Science, Science and Research Branch, Islamic Azad University, Tehran, Iran
Seyed-Hassan Mirian-Hosseinabadi, Department of Computer Science, Sharif University of Technology, Tehran, Iran*

Abstract - Evolutionary software development disciplines, such as Agile Development (AD), are test-centered, and their application in model-based frameworks requires model support for test development. These tests must be applied

against changes during software evolution. Traditionally regression testing exposes the scalability problem, not only in terms of the size of test suites, but also in terms of complexity of the formulating modifications and keeping the fault detection after system evolution. Model Driven Development (MDD) has promised to reduce the complexity of software maintenance activities using the traceable change management and automatic change propagation. In this paper, we propose a formal framework in the context of agile/lightweight MDD to define generic test models, which can be automatically transformed into executable tests for particular testing template models using incremental model transformations. It encourages a rapid and flexible response to change for agile testing foundation. We also introduce on-the-fly agile testing metrics which examine the adequacy of the changed requirement coverage using a new measurable coverage pattern. The Z notation is used for the formal definition of the framework. Finally, to evaluate different aspects of the proposed framework an analysis plan is provided using two experimental case studies.

Keywords: Agile development, Model Driven testing, On-the-fly Regression Testing, Model Transformation, Test Case Selection.

36. Paper 29021681: Comparative Analysis of Early Detection of DDoS Attack and PPS Scheme against DDoS Attack in WSN (pp. 334-342)

*Kanchan Kaushal, Varsha Sahni
Department of Computer Science Engineering, CTIEMT Shahpur Jalandhar, India*

Abstract- Wireless Sensor Networks carry out has great significance in many applications, such as battlefields surveillance, patient health monitoring, traffic control, home automation, environmental observation and building intrusion surveillance. Since WSNs communicate by using radio frequencies therefore the risk of interference is more than with wired networks. If the message to be passed is not in an encrypted form, or is encrypted by using a weak algorithm, the attacker can read it, and it is the compromise to the confidentiality. In this paper we describe the DoS and DDoS attacks in WSNs. Most of the schemes are available for the detection of DDoS attacks in WSNs. But these schemes prevent the attack after the attack has been completely launched which leads to data loss and consumes resources of sensor nodes which are very limited. In this paper a new scheme early detection of DDoS attack in WSN has been introduced for the detection of DDoS attack. It will detect the attack on early stages so that data loss can be prevented and more energy can be reserved after the prevention of attacks. Performance of this scheme has been seen by comparing the technique with the existing profile based protection scheme (PPS) against DDoS attack in WSN on the basis of throughput, packet delivery ratio, number of packets flooded and remaining energy of the network.

Keywords: DoS and DDoS attacks, Network security, WSN

37. Paper 29021687: Detection of Stealthy Denial of Service (S-DoS) Attacks in Wireless Sensor Networks (pp. 343-348)

*Ram Pradheep Manohar, St. Peter's University, Chennai
E. Baburaj, Narayanaguru College of Engineering, Nagercoil*

Abstract — Wireless sensor networks (WSNs) supports and involving various security applications like industrial automation, medical monitoring, homeland security and a variety of military applications. More researches highlight the need of better security for these networks. The new networking protocols account the limited resources available in WSN platforms, but they must tailor security mechanisms to such resource constraints. The existing denial of service (DoS) attacks aims as service denial to targeted legitimate node(s). In particular, this paper address the stealthy denial-of-service (S-DoS) attack, which targets at minimizing their visibility, and at the same time, they can be as harmful as other attacks in resource usage of the wireless sensor networks. The impacts of Stealthy Denial of Service (S-DoS) attacks involve not only the denial of the service, but also the resource maintenance costs in terms of resource usage. Specifically, the longer the detection latency is, the higher the costs to be incurred. Therefore, a particular attention has to be paid for stealthy DoS attacks in WSN. In this paper, we propose a new attack strategy namely Slowly Increasing and Decreasing under Constraint DoS Attack Strategy (SIDCAS) that leverage the application vulnerabilities, in order to degrade the performance of the base station in WSN. Finally we analyses the characteristics of the S-DoS attack against the existing Intrusion Detection System (IDS) running in the base station.

38. Paper 2902169912: Intelligent Radios in the Sea (pp. 349-357)

Ebin K. Thomas

Amrita Center for Wireless Networks & Applications (AmritaWNA), Amrita School of Engineering, Amritapuri, Amrita Vishwa Vidyapeetham, Amrita University, India

Abstract - Communication over the sea has huge importance due to fishing and worldwide trade transportation. Current communication systems around the world are either expensive or use dedicated spectrum, which lead to crowded spectrum usage and eventually low data rates. On the other hand, unused frequency bands of varying bandwidths within the licensed spectrum have led to the development of new radios termed Cognitive radios that can intelligently capture the unused bands opportunistically by sensing the spectrum. In a maritime network where data of different bandwidths need to be sent, such radios could be used for adapting to different data rates. However, there is not much research conducted in implementing cognitive radios to maritime environments. This exploratory article introduces the concept of cognitive radio, the maritime environment, its requirements and surveys, and some of the existing cognitive radio systems applied to maritime environments.

Keywords — *Cognitive Radio, Maritime Network, Spectrum Sensing.*

39. Paper 290216991: Developing Context Ontology using Information Extraction (pp. 358-363)

*Ram kumar #, Shailesh Jaloree #, R S Thakur **

Applied Mathematics & Computer Application, SATI, Vidisha, India

** MANIT, Bhopal, India*

Abstract — Information Extraction addresses the intelligent access to document contents by automatically extracting information applicable to a given task. This paper focuses on how ontologies can be exploited to interpret the contextual document content for IE purposes. It makes use of IE systems from the point of view of IE as a knowledge-based NLP process. It reviews the dissimilar steps of NLP necessary for IE tasks: Rule-Based & Dependency Based Information Extraction, Context Assessment.

40. Paper 31031601: Challenges and Interesting Research Directions in Model Driven Architecture and Data Warehousing: A Survey (pp. 364-398)

Amer Al-Badarneh, Jordan University of Science and Technology

Omran Al-Badarneh, Devoteam, Riyadh, Saudi Arabia

Abstract - Model driven architecture (MDA) is playing a major role in today's system development methodologies. In the last few years, many researchers tried to apply MDA to Data Warehouse Systems (DW). Their focus was on automatic creation of Multidimensional model (Start schema) from Conceptual Models. Furthermore, they addressed the conceptual modeling of QoS parameters such as Security in early stages of system development using MDA concepts. However, there is a room to improve further the DW development using MDA concepts. In this survey we identify critical knowledge gaps in MDA and DWs and make a chart for future research to motivate researchers to close this breach and improve DW solution's quality and performance, and also minimize drawbacks and limitations. We identified promising challenges and potential research areas that need more work on it. Using MDA to handle DW performance, multidimensionality and friendliness aspects, applying MDA to other stages of DW development life cycle such as Extracting, Transformation and Loading (ETL) Stage, developing On Line Analytical Processing(OLAP) end user Application, applying MDA to Spatial and Temporal DWs, developing a complete, self-contained DW framework that handles MDA-technical issues together with managerial issues using Capability Maturity Model Integration(CMMI) standard or International standard Organization (ISO) are parts of our findings.

Keywords: Data warehousing, Model driven Architecture (MDA), Platform Independent Model (PIM), Platform Specific Model (PSM), Common Warehouse Metamodel (CWM), XML Metadata Interchange (XMI).

41. Paper 290216992: A Framework for Building Ontology in Education Domain for Knowledge Representation (pp. 399-403)

Monica Sankat #, R S Thakur, Shailesh Jaloree #*

Department of Applied Mathematics and Computer Application, SATI, Vidisha, India

** Department of Computer Application, MANIT, Bhopal, India*

Abstract — In this paper we have proposed a method of creating domain ontology using protégé tool. Existing ontology does not take the semantic into context while displaying the information about different modules. This paper proposed a methodology for the derivation and implementation of ontology in education domain using protégé 4.3.0 tool.

42. Paper 2902169913: Mobility Aware Multihop Clustering based Safety Message Dissemination in Vehicular Ad-hoc Network (pp. 404-417)

Nishu Gupta, Dr. Arun Prakash, Dr. Rajeev Tripathi

Department of Electronics and Communication Engineering

Motilal Nehru National Institute of Technology Allahabad, Allahabad-211004, INDIA

Abstract - A major challenge in Vehicular Ad-hoc Network (VANET) is to ensure real-time and reliable dissemination of safety messages among vehicles within a highly mobile environment. Due to the inherent characteristics of VANET such as high speed, unstable communication link, geographically constrained topology and varying channel capacity, information transfer becomes challenging. In the multihop scenario, building and maintaining a route under such stringent conditions becomes even more challenging. The effectiveness of traffic safety applications using VANET depends on how efficiently the Medium Access Control (MAC) protocol has been designed. The main challenge while designing such a MAC protocol is to achieve reliable delivery of messages within the time limit under highly unpredictable vehicular density. In this paper, Mobility aware Multihop Clustering based Safety message dissemination MAC Protocol (MMCS-MAC) is proposed in order to accomplish high reliability, low communication overhead and real time delivery of safety messages. The proposed MMCS-MAC is capable of establishing a multihop sequence through clustering approach using Time Division Multiple Access mechanism. The protocol is designed for highway scenario that allows better channel utilization, improves network performance and assures fairness among all the vehicles. Simulation results are presented to verify the effectiveness of the proposed scheme and comparisons are made with the existing IEEE 802.11p standard and other existing MAC protocols. The evaluations are performed in terms of multiple metrics and the results demonstrate the superiority of the MMCS-MAC protocol as compared to other existing protocols related to the proposed work.

Keywords- Clustering, Multihop, Safety, TDMA, V2V, VANET.

43. Paper 290216997: Clustering of Hub and Authority Web Documents for Information Retrieval (pp. 418-422)

Kavita Kanathey, Computer Science, Barkatullah University, Bhopal, India

R. S. Thakur, Department of Computer Application, Maulana Azad National Institute of Technology (MANIT), Bhopal, MP, India

Shailesh Jaloree, Department of applied Mathematics, SATI, Vidisha, Bhopal, MP, India

Abstract - Due to the exponential growth of World Wide Web (or simply the Web), finding and ranking of relevant web documents has become an extremely challenging task. When a user tries to retrieve relevant information of high quality from the Web, then ranking of search results of a user query plays an important role. Ranking provides an ordered list of web documents so that users can easily navigate through the search results and find the information content as per their need. In order to rank these web documents, a lot of ranking algorithms (PageRank, HITS, Weight

PageRank) have been proposed based upon many factors like citations analysis, content similarity, annotations etc. However, the ranking mechanism of these algorithms gives user with a set of non-classified web documents according to their query. In this paper, we propose a link-based clustering approach to cluster search results returned from link based web search engine. By filtering some irrelevant pages, our approach classified relevant web pages into most relevant, relevant and irrelevant groups to facilitate users' accessing and browsing. In order to increase relevancy accuracy, K-mean clustering algorithm is used. Preliminary evaluations are conducted to examine its effectiveness. The results show that clustering on web search results through link analysis is promising. This paper also outlines various page ranking algorithms.

Keywords - World Wide Web, search engine, information retrieval, Pagerank, HITS, Weighted Pagerank, link analysis.

44. Paper 29021645: Dorsal Hand Vein Identification (pp. 423-433)

Sarah HACHEMI BENZIANE, Irit, University Paul Sabatier France, Simpa, University of Sciences and Technology of Oran, Mohamed Boudiaf Algérie

Abdelkader BENYETTOU, Simpa, University of Sciences and Technology of Oran

Abstract — In this paper, we present an competent approach for dorsal hand vein features extraction from near infrared images. The physiological features characterize the dorsal venous network of the hand. These networks are single to each individual and can be used as a biometric system for person identification/authentication. An active near infrared method is used for image acquisition. The dorsal hand vein biometric system developed has a main objective and specific targets; to get an electronic signature using a secure signature device. In this paper, we present our signature device with its different aims; respectively: The extraction of the dorsal veins from the images that were acquired through an infrared device. For each identification, we need the representation of the veins in the form of shape descriptors, which are invariant to translation, rotation and scaling; this extracted descriptor vector is the input of the matching step. The optimization decision system settings match the choice of threshold that allows to accept / reject a person, and selection of the most relevant descriptors, to minimize both FAR and FRR errors. The final decision for identification based descriptors selected by the PSO hybrid binary give a FAR =0% and FRR=0% as results.

Keywords - Biometrics, identification, hand vein, OTSU, anisotropic diffusion filter, top & bottom hat transform, BPSO,

45. Paper 290216999: Blind Image Separation Based on Exponentiated Transmuted Weibull Distribution (pp. 423-433)

A. M. Adam, Department of Mathematics, Faculty of Science, Zagazig University, P.O. Box, Zagazig, Egypt

R. M. Farouk, Department of Mathematics, Faculty of Science, Zagazig University, P.O. Box, Zagazig, Egypt

M. E. Abd El-aziz, Department of Mathematics, Faculty of Science, Zagazig University, P.O. Box, Zagazig, Egypt

Abstract - In recent years the processing of blind image separation has been investigated. As a result, a number of feature extraction algorithms for direct application of such image structures have been developed. For example, separation of mixed fingerprints found in any crime scene, in which a mixture of two or more fingerprints may be obtained, for identification, we have to separate them. In this paper, we have proposed a new technique for separating a multiple mixed images based on exponentiated transmuted Weibull distribution. To adaptively estimate the parameters of such score functions, an efficient method based on maximum likelihood and genetic algorithm will be used. We also calculate the accuracy of this proposed distribution and compare the algorithmic performance using the efficient approach with other previous generalized distributions. We find from the numerical results that the proposed distribution has flexibility and an efficient result.

Keywords- Blind image separation, Exponentiated transmuted Weibull distribution, Maximum likelihood, Genetic algorithm, Source separation, FastICA.

PSNR and Jitter Analysis of Routing Protocols for Video Streaming in Sparse MANET Networks, using NS2 and the Evalvid Framework

Sabrina Nefti

Dept. of Computer Science
University Batna 2
Algeria

Maamar Sedrati

Dept. of Computer Science
University Batna 2
Algeria

Abstract—Advances in multimedia and ad-hoc networking have urged a wealth of research in multimedia delivery over ad-hoc networks. This comes as no surprise, as those networks are versatile and beneficial to a plethora of applications where the use of fully wired network has proved intricate if not impossible, such as prompt formation of networks during conferences, disaster relief in case of flood and earthquake, and also in war activities. In this paper, we aim to investigate the combined impact of network sparsity and network node density on the Peak Signal Noise to Ratio (PSNR) and jitter performance of proactive and reactive routing protocols in ad-hoc networks. We also shed light onto the combined effect of mobility and sparsity on the performance of these protocols. We validate our results through the use of an integrated Simulator-Evaluator environment consisting of the Network Simulator NS2, and the Video Evaluation Framework Evalvid.

Keywords- PSNR, MANET, Sparsity, Density, Routing protocols, Video Streaming, NS2, Evalvid

I. INTRODUCTION

The transmission of multimedia objects over Mobile Ad-hoc Networks (MANET) network has become the need of the day due to critical applications that rely on such networks such as the transmission of important images and videos in emergency situations. However, this task presents two main complexities. The first aspect of intricacy lies in the nature of MANET: their mobile and distributed, interference and multi-hop communication [1], [2]. Since MANETs do not rely on pre-existing infrastructure, data is transmitted through multi-hop routing [3]. This collective effort in data transmission requires that each node acts as a router too. Thus, it comes as no surprise that the provision of QoS over such networks can prove extremely difficult. The second aspect of intricacy lies within the nature of multimedia objects, specifically video files which are not only bandwidth-hungry but also highly-demanding in terms of Quality of Service (QoS). Effective multimedia transmission dictates minimal delay and in-order receipt of packets [1]. Therefore, it has become imperative to

determine routing protocols that can not only fulfil those QoS criteria but are also able to maintain such performance while varying the network topology in terms of sparsity and mobility. Multimedia transmission may prove particularly challenging in sparse MANETs whereby disconnections become more and more frequent due to low network node density [3].

In this paper, we investigate the PSNR performance of proactive and reactive routing protocols for video streaming of bandwidth-hungry multimedia video files over sparse MANET networks. We also explore the combined effect of mobility and sparsity on the PSNR performance. To this end, we use the NS2 Network Simulator and the Evalvid Framework tool in order to test a renowned protocol of each family, namely AODV (reactive), and DSDV (proactive).

The remainder of the paper is organized as follows: Section 2 explores the previous work performed in this field. A brief description of the System Model adopted in our work is presented in Section 3. Section 4 justifies our choice of the simulation and evaluation tools used. Next, a detailed work approach is described along with simulation configuration in Section 5. Results are presented and analyzed in Section 6. Finally, conclusions and future work recommendations are provided in Section 7.

II. RELATED WORK

Various comparative studies have been carried out between proactive and reactive protocols [4], [5]. In [5], the QoS metrics used for comparison are media access delay, network load and throughput. The study in [6] is similar to [5] with the addition of retransmission attempts metric. However, the aforementioned studies have not taken into consideration the augmented challenges dictated by the transmission of quality-demanding multimedia objects. A more specific analysis of routing protocols for video streaming was undertaken in [7], whereby two network structures (25 nodes and 81 nodes) were simulated using OPNET in order to assess QoS parameters such as throughput, wireless LAN delay, end-to-end delay and packet delay variation. In our work, we investigate the performance of MANET routing protocols in video streaming

on the basis of Peak to Signal Noise Ratio Video Quality Model, also called VQM_p [8] as well as jitter.

III. SYSTEM MODEL

Fig. 1 shows the overall System Model: initially a video file in the raw YUV format is converted into a MPEG4 file. This latter is fed into Evalvid to generate a video trace file which is the actual video object that is sent over a simulated transmission in NS2. The received video file received at the receiver node in the simulation environment is fed back into Evalvid which generates the PSNR quality model amongst various other QoS metrics [9].

The PSNR quality model is considered as one of the most widespread models in assessing video quality in an objective manner, as it was developed specifically to emulate the quality impression of the Human Visual System (HVS) [11]. Furthermore, this model is a derivative of the notorious Signal to Noise Ratio (SNR). However, while SNR compares the signal energy to the error energy, PSNR compares the maximum possible signal energy to the noise energy. This subtle difference has shown to yield higher correlation with the subjective quality perception than the conventional SNR [12]. The following equation is the definition of the PSNR between the luminance component Y of source image S and destination image D [11]:

$$PSNR(n)_{dB} = 20 \log_{10} \left(\frac{V_{peak}}{\sqrt{\frac{1}{N_{col} N_{row}} \sum_{i=0}^{N_{col}} \sum_{j=0}^{N_{row}} [Y_s(n, i, j) - Y_D(n, i, j)]^2}} \right) \quad (1)$$

$$V_{peak} = 2^k - 1 \quad k = \text{number of bits per pixel} \quad (2)$$

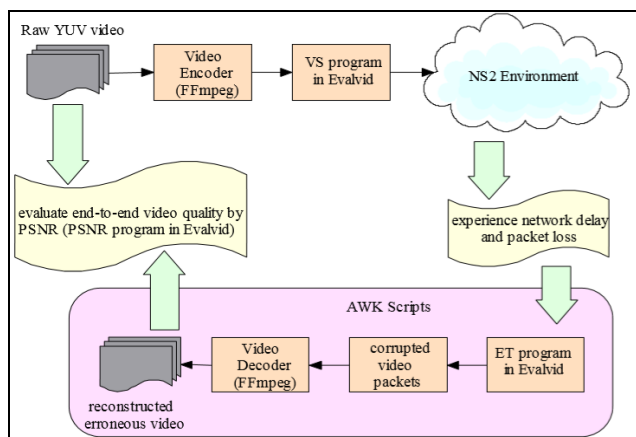


Fig. 1: System Model: Evalvid Framework integration with NS2 [10]

IV. SIMULATION AND EVALUATION TOOLS

A survey of simulation tools for MANET has revealed that the NS2 simulator is one of the most widespread tools for MANET network simulations with a usage percentage of 43.8%, largely outperforming its direct competitor namely the bespoke tools (27.3%), as can be depicted in Fig. 2 [13]. Based on this finding, we opted for NS2 in our work.

Two main video quality evaluation tools were identified in the literature, namely the MSU tool used to extract the Structural SIMilarity (SSIM) index [14] and the Evalvid Framework, which evaluates the quality of videos transmitted over real networks [11]. The advantage that Evalvid presents compared to MSU is that it can be integrated into NS2, and therefore it is possible to use Evalvid to evaluate the quality of a video object transmitted in an NS2 simulation environment [12], hence the reason for which we selected Evalvid Framework in our work.

A. The Evalvid Tools

The Evalvid Framework consists of three main tools [10]. In the following tables, we summarize the functionality and the parameters of these tools based on their execution in the Windows command line tool (cmd).

TABLE 1: FUNCTIONALITY AND PARAMETERS OF THE “MP4TRACE” TOOL

Tool	mp4trace
Functionality	Converts the MPEG4 video file to be transmitted into a video trace file. This trace file is then sent by the source node in the NS2 simulation environment to the receiver node.
Usage	<code>mp4trace [options] <file 1> <file 2></code>
Parameters	Options: -[p/f] packet or frame mode -s host port: sends the RTP packets to specified host and UDP port <file 1> the MPEG4 video file to be transmitted <file 2> the generated trace file generated by the tool

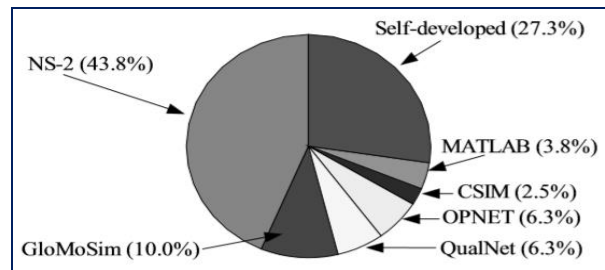


Fig. 2: Utilization percentages of Simulation Software for MANET [13]

TABLE 2: FUNCTIONALITY AND PARAMETERS OF THE “ETMP4” TOOL

Tool	etmp4
Functionality	1. The video trace file “st” generated by <i>mp4trace</i> (Table1) is fed into NS2 and two files are generated from the simulation: the sender’s frame transmission times file “sd”, and the

	<p>receiver's frame reception times file "rd". The <i>etmp4</i> function generates the mpeg4 video file "out" that was received at the receiver based on the "st", "sd" and "rd" files as well as the original transmitted video file.</p> <p>2. This tool also generates QoS metrics such as loss rate, debit etc...</p>
Usage	etmp4 -[p f F] -[0 x] [-c] <sd> <rd> <st> <in> <out>
Parameters	<p>-[p f F] packet, frame or complete frame mode</p> <p>-[0 x] fill lost section with 0 or truncate</p> <p>[-c] use cumulative jitter in case of asynchronous clocks</p> <p><sd> tcpdump sender</p> <p><rd> tcpdump receiver</p> <p><st> trace-file sender</p> <p><in> transmitted video (original mp4)</p> <p><out> base name of output file</p> <p>[PoB] optional Play-out buffer size [ms]</p>

TABLE 3: FUNCTIONALITY AND PARAMETERS OF THE "PSNR" TOOL

Tool	Psnr
Functionality	This function generates the PSNR metric image by image according to the aforementioned formulae.
Usage	psnr x y <YUV format> <src.yuv> <dst.yuv>
Parameters	<p>x frame width</p> <p>y frame height</p> <p>YUV format: 420, 422, etc..</p> <p>src.yuv: source video</p> <p>dst.yuv: distorted video</p>

B. The Integration of Evalvid into NS2

In order to integrate the Evalvid Tool in NS2, two main types of amendments are required:

- Modifications applied in the NS2 header files and Makefile, these are explained in detail in [10].
- The addition of new C++ classes into the NS2 core code [15][10]: Those classes augment the NS2 environment with objects that have the capability to transmit video trace files (generated by the *mp4trace* tool described in Table 1). We analyzed the code of these classes and presented our understanding of their augmented video capabilities in Table 4.

TABLE 4: THE CONTRIBUTION OF THE NEW CLASSES

New Class	The Contribution of the New Class in the Video Streaming Simulation
myUDP	<p><i>myUDP</i> class extends the NS2 Agent Class (and hence has access to the latter's functions, thanks to the inheritance principle). An object of <i>myUDP</i> represents the UDP transport protocol and has two main video capabilities:</p> <ol style="list-style-type: none"> 1. The function «<i>attach-agent</i>», attaches the <i>myUDP</i> object to a <i>myEvalvid</i> object. Thanks to this attachment, the <i>myUDP</i> object can extract the video trace file (to be transmitted) from the <i>myEvalvid</i> object (to which it is attached). 2. The function "<i>set_filename</i>", passes to

	<p>the object <i>myUDP</i> a file pointer in which it can record the transmission time of each transmitted frame (or packet).</p>
MyEvalvid_Sink	<p>Similar to the above, the <i>MyEvalvid_Sink</i> class extends the NS2 Agent Class. The main two functionalities that contribute to the video transmission process over NS2 are as follows:</p> <ol style="list-style-type: none"> 1. The function «<i>connect</i>» connects the <i>myUDP</i> object to the <i>MyEvalvid_Sink</i> object. This connection renders possible the transmission of the video trace file (attached to the <i>myUDP</i> sender object) to the receiver object <i>MyEvalvid_Sink</i>. 2. The function «<i>set_filename</i>» allows the <i>MyEvalvid_Sink</i> object to record the reception time of each received frame (or packet).
myEvalvid	<p><i>myEvalvid</i> Class extends the NS2 Traffic Class. An object of type <i>myEvalvid</i> represents a traffic source of type video and has two main capabilities:</p> <ol style="list-style-type: none"> 1. Thanks to the function «<i>attach-tracefile</i>», the <i>myEvalvid</i> object can be attached to an object of type <i>Trace</i> (an inherent NS2 Class). The <i>Trace</i> object, in turn, can be attached to a video trace file (generated by <i>mp4trace</i> tool) thanks to the function "<i>filename</i>". This double attachment enables the <i>myEvalvid</i> object to be attached to the video trace file (to be transmitted). 2. Thanks to the function «<i>attach-agent</i>», the <i>myUDP</i> object can be attached to the traffic source object <i>myEvalvid</i> which encompasses the video trace file (see point 1 above). This attachment enables the transport object <i>myUDP</i> to transport the video trace file (as this data can be obtained from the traffic source object <i>myEvalvid</i> attached to the transport object <i>myUDP</i>).

V. WORK APPROACH

Our work approach is summarized in four main steps as shown in Fig. 3. Each step is explained separately in the following sub-sections.

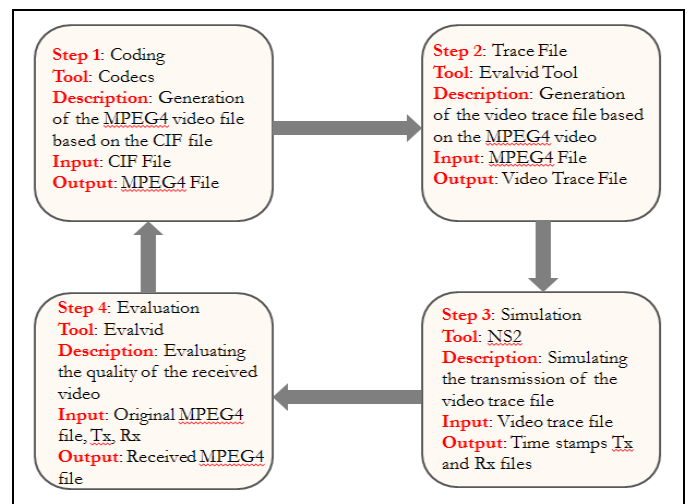


Fig. 3: Work Steps: (1) Coding (2) Trace File Generation (3) NS2 Simulation (4) Evaluation in Evalvid

A. The Coding Process

The type of video files used in our simulations are the H.261 standard Common Interface Format (CIF) with 352 × 288 resolution, as this format is commonly used in video teleconferencing, which is one of the important applications of MANET. Before a CIF file could be used for simulation purposes, it is first encoded into MPEG format, one of industrial standards widely used in video streaming over the internet [16]. This coding occurs in three stages [10]:

- First the CIF file is converted into a YUV file, a video data format which takes into account human perception. This is performed by the ffmpeg codec [17]. The command line is as follows:

```
ffmpeg -i CIF_File Original_Yuv_File.yuv
```

- The Yuv file resulting from the above is converted MP4V format with the xvi codec [17]. This format is considered as the intermediary raw format of MP4:

```
xvid_encraw -i Original_Yuv_File.yuv -w 352 -h 288 -framerate 30 -max_key_interval 30 -o Original_MV4_File.m4v
```

- Finally, the M4V file (*Original_MV4_File.m4v*) is coded into a MPEG4 file (*Original_MP4_File.mp4*) with the MP4Box codec [17], with the following command line:

```
MP4Box -hint -mtu 1024 -fps 30 -add Original_MV4_File.m4v Original_MP4_File.mp4
```

B. Video Trace File Generation

The second step consists of generating a video trace file from the original MPEG4 video file using the *mp4trace* tool described in Section 4.1. The video trace file contains the frame number, type and size and the number of segments in case of frame segmentation [11]. The Evalvid tool was originally designed to evaluate real video transmissions, hence the reason why *mp4trace* tool specifies the destination URL and port number. However, for the sake of our work, this tool is executed with an arbitrary IP and Port number, as the sole aim of this execution is the generation of the video trace file and not its actual transmission over the internet:

```
mp4trace -f -s 192.168.0.2 12346 Original_MP4_File .mp4 > eval_trace_file
```

C. Simulation in NS2

In the first set of simulations, we aim to investigate the combined impact of the network node density and network sparsity on the PSNR performance of AODV and DSDV. To this end, we designed a matrix-based network topology whereby nodes are symmetrically placed in a matrix with equal horizontal and vertical distances from each other. By network sparsity, we mean how distanced or close the network nodes are, whereby the distance between each neighboring nodes in the matrix topology is referred to by Distance (D). By network node density, we mean the number of nodes that the network consists of.

We begin our simulation with a network with nodes that are close to each other (as opposed to sparse network), and gradually disperse it, by equally augmenting the vertical and horizontal distance between each two neighboring nodes. The distances considered are 20m, 50m, 100m, and 150m. In order to test the combined effect of network node density and its sparsity, we test each network sparsity model with several network nodes density ranging from 4, 9, 16, 25, 36, 49 and 64 nodes. Table 5 shows our simulation configuration.

In the second set of simulations, we test the impact of mobility that results in sparsity on the PSNR performance of both AODV and DSDV. To this end, we start the simulation with a network in which the nodes are closely distanced from each other (D=20m), and which move outward with a constant speed; in order to form a sparse matrix (D=150m), as can be perceived in Fig. 4. We refer to this scenario in the remainder of our paper as *Outward Mobility*.

In the third and last set of simulations, we test the effect of mobility that results in a network with closely distanced nodes on the PSNR performance of both protocols. To this end, we commence the simulation with a sparse network (D=150 m), in which nodes move inward with a constant speed to form a network with closely distanced nodes (D=20m). We refer to this scenario in the rest of our paper as *Inward Mobility*.

TABLE 5: SIMULATION CONFIGURATION

Simulation Parameter	Configuration
Propagation Model	TwoRayGround
MAC	802.11
Routing Protocols	AODV, DSDV
Placement of Nodes	Matrix-based placement with equal vertical and horizontal distance between nodes. This distance varies from: 20m, 50m, 100m, and 150m
Number of Node	4, 9, 16, 25, 49, and 64 Arranged in matrices of: 2 x 2, 3 x 3, 4 x 4, 5 x 5, 6 x 6, 7 x 7, and 8 x 8
Video File Frame Size	2000 frames

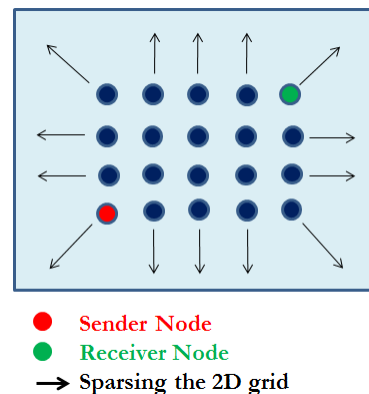


Fig. 4: Outward Mobility: The nodes are dispersed as a result of mobility.

D. Evaluation of the Video Quality

We evaluate of the quality of the received video file in three stages [10] [11]:

- The generation of the received video file: The Evalvid tool `etmp4` compares the sender's timestamp file (which contains the transmission time of each transmitted frame) against the receiver's timestamp file (which contains the reception time of each received frame). Through this comparison, and the original MP4 video and the video trace file, the tool reconstructs the received MP4 video. During this reconstruction process, the tool also measures the delay per frame, the frame loss rate, the instantaneous transmission as well as well as the reception debit. The corresponding command is as follows:

```
etmp4 -f -0 <s_time_trace> <r_time_trace>
<video_trace_file> Original_Mp4_File.mp4
Received_MP4_File.mp4
```

- The generation of the received Yuv file: using the `ffmpeg` codec and the received video (from the previous step). The required command is as follows:

```
ffmpeg -i Received_MP4_File.mp4
Received_Yuv_File.yuv
```

- The generation of the PSNR metric: using the `psnr` Evalvid tool which compares the original and the received Yuv files in order to calculate the PSNR per frame. The corresponding command is as follows:

```
psnr 352 288 420 Original_Yuv_File.yuv
Received_Yuv_File.yuv
```

VI. RESULTS AND ANALYSIS

A. PSNR Performance Analysis

The video used in all our simulations is based on “Highway CIF” [18]. Fig. 5 shows the PSNR performance of AODV when the network sparsity is fixed to $D=100m$ and its density is increased. A moving average filter of a 100 frames width was used to smooth the results a clearer analysis.

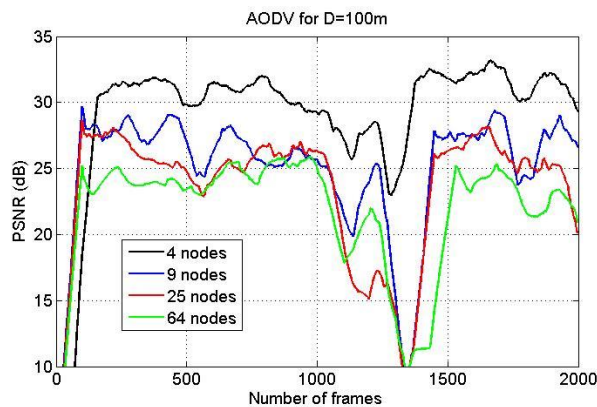


Fig. 5: PSNR Performance of AODV for various network densities (4, 9, 25, 64) when the distance is fixed to $D=100$.

The PSNR variation pattern is maintained across various network densities. The PSNR variation pattern is affected by the luminosity content of the frames that the transmitted video consists of. Since we are transmitting the same video file across varying topologies, the luminosity content of each frame of this video remains constant. We further verified this by investigating the two main drops in PSNR performance in Fig. 5. The first one occurred approximately at frame $F=550$, corresponding to the approximate video play time of $T=21s$. We observed that during this timeframe, the luminosity is decreased due to the appearance of an overtaking black car, as can be depicted in Fig. 6.

The second PSNR major drop occurred between frames $F=1250$ and $F=1300$, corresponding to the approximate video play times of $T=41s$ and $T=43s$. During this timeframe a dark bridge first appears in the video and then the car passes under its shadow as shown in Fig. 7.

Therefore, the PSNR drop in the two cases of Fig. 6 and Fig. 7 can be justified by the fact that when the luminosity content of a frame decreases, the noise energy dominates over the peak signal energy, and hence degrading the PSNR.

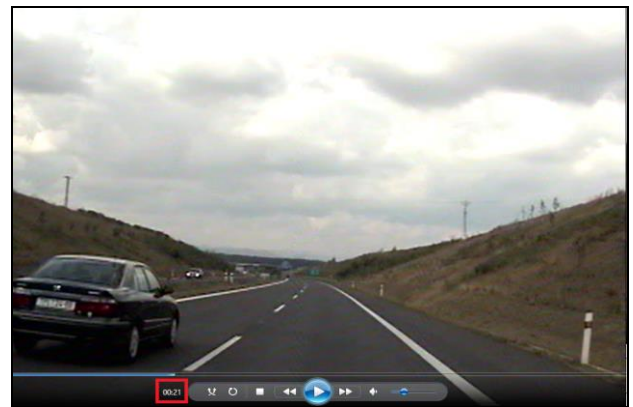


Fig. 6: Appearance of a black car at $T=21s$



Fig. 7: Appearance of a bridge and its shadow at $T=41s$

It can also be noted from Fig. 5 that the PSNR performance of AODV degrades as the density of the network increases. The PSNR decreases by approximately 5dB between frames Frame=400 and Frame=600 when upsizing the network from N=4 nodes to N=9 nodes. However, this attenuation is less significant when the network is further upsized to N=25 nodes and N=64 nodes. This is due to the fact that the 4-nodes topology allows direct communication between the sender and the receiver nodes; however, as the network density is augmented to N=9 nodes, data is routed through intermediary nodes. In this case, multi-hopping decreases the PSNR performance significantly compared to direct sender receiver one-hop transmission.

It can also be seen that AODV registers a sharp rise in PSNR between frame F=0 and F=100 for the various network topologies, which can be justified by its fast convergence in low density networks [19].

Fig. 8 compares the PSNR performance of AODV and DSDV in two different topologies. A moving average filter of a 100 frames width was used to smooth the results for a clearer analysis. It is clear that AODV outperforms DSDV in small sparse networks (D=100m, N=4), with a PSNR difference ranging from 5dB to 23dB. When the distance between two neighboring nodes is halved from D = 100m to D=50m, the PSNR performance of both protocols increases by a range of 3dB to 10dB and also becomes smoother. This occurs despite the fact that the network density is quadrupled from N=4 to N=16 nodes.

In this scenario, it can be concluded that PSNR performance of both AODV and DSDV is better in high-density and low-sparsity networks than in it low-density and high-sparsity ones. In fact, extensive simulations demonstrated that when increasing the network sparsity, DSDV protocol is unable to deliver a video quality that is sufficient enough to extract the PSNR metric; Table 6 captures these cases.

In order to verify this finding, we reconstructed and played the video sent over a network topology of N = 64 nodes and D=100m in which DSDV was set as the routing protocol, and noted that it was significantly distorted for more than half of the video length as captured in Fig. 9.

B. Jitter Performance Analysis

Figure 10 shows the jitter performance of ADOV when varying the network density. It can be remarked that the jitter variation pattern remains similar when varying network densities.

Similar to PSNR, the jitter variation pattern is also dependent on the luminosity content of the frames transmitted. However, unlike PSNR, jitter performance improves when the frame luminosity content is low; as such frames carry less data content than frames with high luminosity, and hence enjoy better delay variation performance. Indeed, the jitter performance improves at approximate frames F=550 and F=1250 which have low luminosity content as discussed earlier, and shown in Figure 6 and Figure 7.

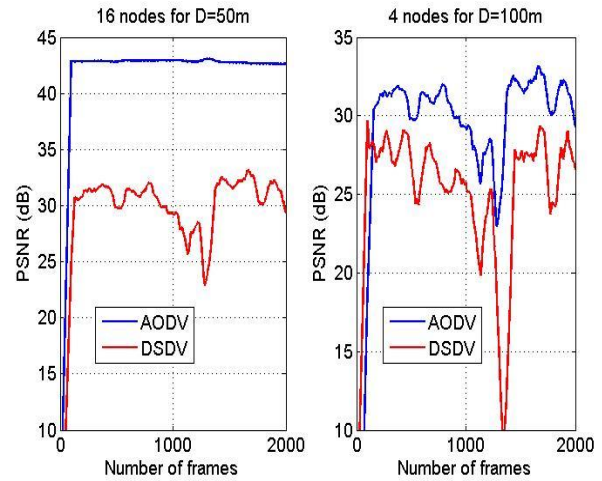


Fig. 8: Comparison of PSNR performance of AODV and DSDV: for two scenarios (D=50m, Nodes=16) and (D=100m, Nodes =4).

TABLE 6: CASES WHERE THE PSNR GENERATION WAS NOT POSSIBLE DUE TO INCREASED NETWORK SPARSITY (A: AODV, D: DSDV, Y: PSNR GENERATED, N: PSNR NOT GENERATED)

No of Nodes	D=20		D= 50		D=100		D=150	
	A	D	A	D	A	D	A	D
4	Y	Y	Y	Y	Y	Y	Y	Y
9	Y	Y	Y	Y	Y	N	Y	N
16	Y	Y	Y	Y	Y	N	Y	N
25	Y	Y	Y	N	Y	N	Y	N
36	Y	Y	Y	N	Y	N	Y	N
49	Y	Y	Y	N	Y	N	Y	N
64	Y	Y	Y	N	Y	N	Y	N



Fig. 9: screen captures of the video received across a network topology of D = 100m, N=64 nodes, DSDV protocol

Fig. 10 also reveals that as we upsize the network, the jitter value is slightly degraded with the worst jitter being registered for N=64 nodes and the overall jitter ranging from approximately -0.1s to 0.05s. This suggests that AODV's jitter performance shows a certain degree of resilience to increasing the network density.

Fig. 11 depicts the jitter performance of DSDV for the same network topologies as Fig. 10. As the network density is augmented, the jitter metric is significantly degraded, with the jitter ranging from -1.5s to -0.2s. Therefore, the DSDV jitter performance shows less resilience to network up-sizing than AODV. This degradation is due to the delay variation incurred by the additional multi-hopping that takes place when the number of intermediary nodes through which data has to be routed increases.

In order to analyze the jitter performance more closely, we considered the various network topologies in Fig. 12. It is clear that AODV outperforms DSDV in all four scenarios, what is interesting to note though, is that for AODV, jitter varies between a small range -0.05s and 0s even when augmenting both the network density and sparsity (from N=16 to N=49 and from D=20m to D=100m). This suggests that AODV's jitter performance is robust to variations in both network density and sparsity.

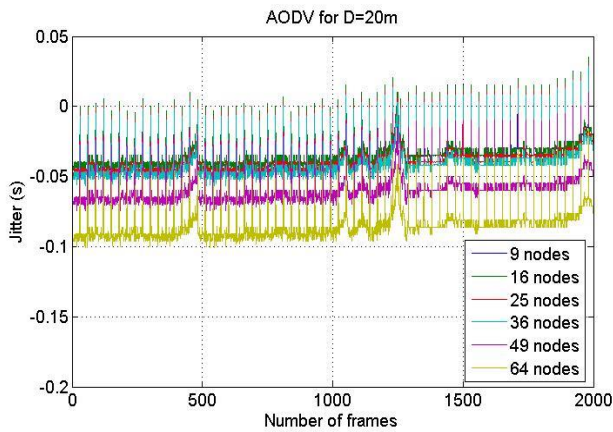


Fig. 10: Jitter performance of AODV with varying network densities (D=20m)

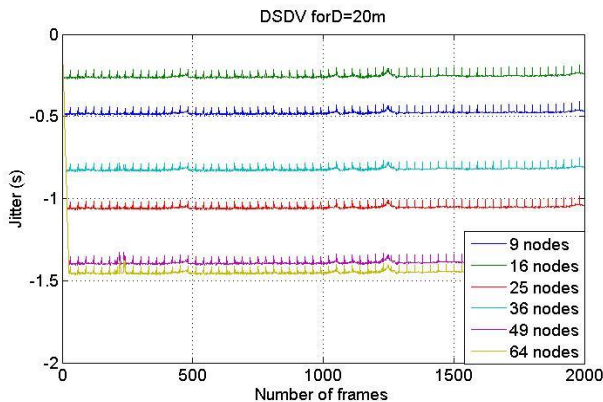


Fig. 11: Jitter performance of DSDV protocol with varying network densities (D=20m)

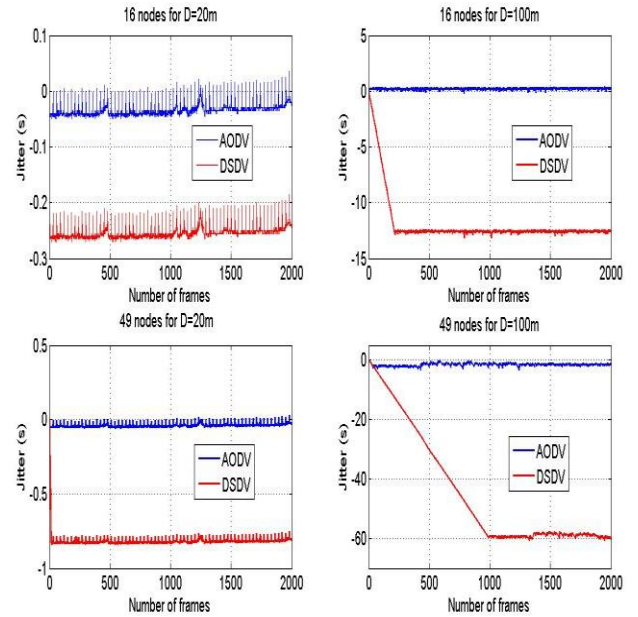


Fig. 12: Comparison of jitter performance for AODV and DSDV for various topologies

DSDV's jitter performance dropped from -0.25s to -0.8s when maintaining the network sparsity (D=20m) and increasing its nodes density (from N=16 to N=49). However, the DSDV's jitter degraded significantly from -0.25s to -12s when the network density was fixed to N=16 and the network sparsity was augmented from D=20m to D=100m. Furthermore, when the density was fixed to N=49, an even more considerable drop in DSDV's jitter performance occurred (from -0.8s to -60s) when the sparsity increased from D=20m to D=100m.

This suggests that AODV's jitter performance shows a much better resilience to the increase in network sparsity and density than DSDV. Furthermore, DSDV's jitter performance is much more resilient to network density than it is to network sparsity, as the results showed that increasing the sparsity of the network degraded the jitter performance more significantly than increasing its density. It can also be noted that as the sparsity increased from D=20m to D=100m, the DSDV's jitter falls sharply in the initial F=250 frames and F=1000 frames respectively, this is due to the DSDV's slow convergence in low density networks [19] as well as increasing the network sparsity.

C. Mobility: PSNR Performance Analysis

Fig. 13 depicts the PSNR performance of a 5x5 matrix in the *Outward Mobility* scenario, i.e. the network's initial sparsity is set to D=20m, and it increases as the nodes move outward with a constant speed to eventually form a sparse matrix of D=150m.

It can be observed that AODV's PSNR performance significantly outperforms DSDV's during the first F=800 frames. This is due to DSDV's slow convergence caused by

periodic updates of routing tables. This is combined with the fact that the network is being dispersed by the nodes movement and hence the update of routing tables between two one-hop neighbours takes longer as the neighbouring nodes are moving away from each other.

However, from frame F=800 onwards, DSDV registers close levels of PSNR compared to AODV. What is interesting to note, is that as the sparsity increases toward the end of the simulation, DSDV's PSNR performance outperforms AODV's. As the mobile nodes become more distant from each other, AODV's reactive route discovery process becomes less efficient, hence reducing the signal strength of each frame compared to its noise content. However, for DSDV, once the routing tables' updates have taken place, DSDV's pre-calculated routes allow for a faster route discovery, hence why DSDV demonstrates a better PSNR performance towards the end of the simulation despite the increasing sparsity of the network.

The opposite scenario is where the network's initial sparsity is set to D=150m and it gradually shrinks to D=20m with the nodes moving inward with a constant speed. Due to the fact that DSDV PSNR metric cannot be extracted when the network is initially very sparse (Table 6), we could not analyse DSDV's PSNR performance in the *Inward Mobility* scenario.

Fig. 14 compares AODV's PSNR performance in *Outward Mobility* against *Inward Mobility*. It can be observed that for approximately the initial 1300 frame, *Outward Mobility* outperforms *Inward Mobility* in terms of PSNR metric. In *Outward Mobility*, PSNR is stronger initially as the nodes are closely positioned from each other, and gradually decreases as the network becomes sparser. This explains the reason why in *Inward Mobility*, the PSNR is weaker initially but eventually outperforms the *Outward Mobility*, as the nodes get closer to each other towards the end of the simulation. Hence, it can be concluded that not only PSNR performance is affected by mobility but is also sensitive to the effect that this mobility has on the network, i.e. whether the mobility results in a sparse network or a closely populated one.

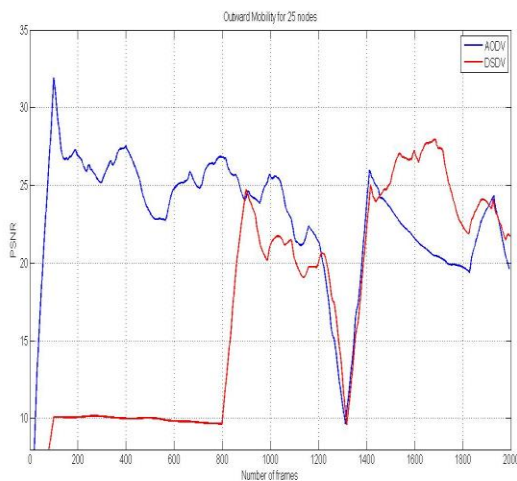


Fig. 13: The impact of Outward Mobility on PSNR for AODV and DSDV (N=25 nodes)

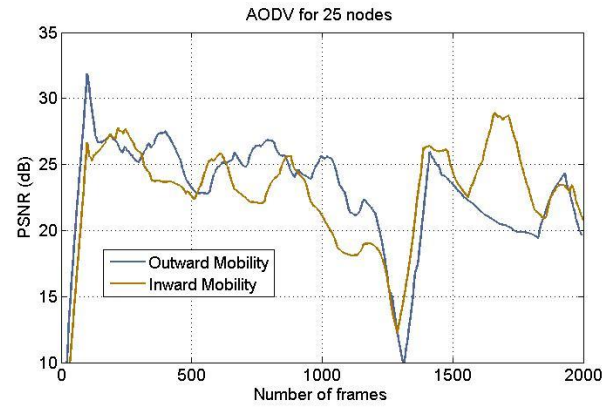


Fig. 14: PSNR performance comparison between inward and outward mobility for AODV

VII. CONCLUSIONS AND PERSPECTIVES

In this paper, we investigated the combined effect of network sparsity and density on PSNR and jitter performances of two MANET routing protocols namely ADOV (reactive) and DSDV (proactive) for video streaming applications. Various network sparsity models were designed with varying network densities. Simulation and evaluation results presented interesting findings. PSNR performance worsens as the density of the network increases. Overall, AODV delivers high levels of PSNR in faster timeframe than DSDV in the various network density and sparsity models analysed in our work. Interestingly, sparsity adversely affects PSNR in a much larger scale than network density for DSDV. In fact, extensive simulations demonstrated that when increasing the network sparsity, DSDV protocol is unable to deliver a video quality that is sufficient enough to extract the PSNR metric.

We also explored the effect two types of mobility on the PSNR metric namely the *Inward* and *Outward Mobility*. It was identified that mobility that results in closely populated networks improves PSNR. Moreover, mobility that results in sparser networks gradually worsens the PSNR performance as the network becomes sparser. Interestingly, in the case of *Outward Mobility*, it was noted that while AODV delivers better PSNR than DSDV initially, DSDV's PSNR outperforms AODV's as the mobile network becomes sparser.

With respect to jitter performance, results demonstrated that AODV's jitter performance is more resilient to changes in both network density and sparsity than DSDV. However DSDV's jitter performance is much more resilient to the increase in network density than it is to the augmentation in network sparsity.

Finally, it is important to highlight that our work relied on a two-dimensional QoS framework, namely PSNR and jitter given that the latter metrics are of paramount importance in video streaming. For future work, we propose undertaking a multi-dimensional QoS comparative study that encompasses the following QoS metrics: frame loss rate, transit delay, throughput, in addition to PSNR and jitter. Such study will allow a closer analysis of the impact of network density,

sparsity, and mobility. We also propose that such a study covers further examples of reactive, proactive and hybrid protocols such as: TORA, OLSR and ZRP.

REFERENCES

- [1] S. Ahmad, and J. Reddy, " Delay optimization using Knapsack algorithm for multimedia traffic over MANETs ", in *Expert Systems with Applications*, 2015, vol 42, no 20, pp. 6819-6827.
- [2] A. Jamali, and N. Naja, " Comparative analysis of ad hoc networks routing protocols for multimedia streaming ", in *Multimedia Computing and Systems ICMCS'09, IEEE International Conference*, 2009, pp. 381-385.
- [3] M. Rao, and N. Singh, "Performance analysis of AODV nthBR protocol for multimedia transmission under different traffic conditions for sparse and densely populated MANETs", in *IEEE Green Computing and Internet of Things (ICGCIoT)*, IEEE International Conference, 2015, pp. 1010-1015.
- [4] P. Sakalley, J. Kumar, " Review and Analysis of Various Mobile Ad Hoc Network Routing Protocols ", *International Journal of Recent Technology and Engineering (IJRTE) ISSN*, 2013, pp.2277-3878.
- [5] P. K. Bhardwaj, and S. Sharma, and V. Dubey, " Comparative analysis of reactive and proactive protocol of mobile ad-hoc network ", *International Journal on Computer Science and Engineering*, Vol. 4, No. 7, 2012, pp. 1281.
- [6] J. Singh, and U. Goyal, " An Analysis of Ad hoc Routing Protocols ", *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, Vol. 1, No. 2, 2015.
- [7] M. Riaz, M. S. I. M. Adnan, and M. Tariqu, " Performance analysis of the Routing protocols for video Streaming over mobile ad hoc Networks ", *International Journal of Computer Networks & Communications*, Vol. 4, No. 3, 2012, pp. 133-150.
- [8] S. Wolf, and M. Pinson, "Video quality measurement techniques", Technical Report 02-392, Department of Commerce, NTIA, USA, 2002.
- [9] C. H. Ke, C. K. Shieh, W. Hwang, and A. Ziviani, " An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission ", *J. Inf. Sci. Eng.*, Vol. 24, No. 2, 2008, pp. 425-440.
- [10] C. H. Ke, "How to evaluate MPEG video transmission using the NS2 simulator?", in <http://csie.nqu.edu.tw/smallko>, EE Department, NCKU, Taiwan University, Taiwan, 2006.
- [11] J. Klaue, J. Rathke, and A. Wolisz, "Evalvid–A framework for video transmission and quality evaluation", in *Computer performance evaluation, Modelling techniques and tools*, Springer Berlin Heidelberg, 2003, pp. 255-272.
- [12] L. Hanzo, and P. J. Cherriman, *Wireless Video Communications*, 445 Hoes Lane, Piscataway, 200: IEEE Press, 2001.
- [13] K. Stuart, C. Tracy, and C. Michael, " MANET simulation studies: the incredible ", in *ACM Mobile Comput Comm Rev*, 2005, Vol. 9, No 4.
- [14] E. Aguiar, A. Riker, A. Abelém, A. Cerqueira, and M. Mu, " Video quality estimator for wireless mesh networks ", in *Quality of Service IEEE 20th International Workshop (IWQoS)*, 2012, pp. 1-9.
- [15] K. Chih-Heng, C. Shieh, W. Hwang, and A. Ziviani, " An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission ", *J. Inf. Sci.*, Vol. 24, No. 2, 2008, pp. 425-440.
- [16] I. Agi, and L. Gong, " An empirical study of secure MPEG video transmissions " in *Network and Distributed System Security, Proceedings of the IEEE Symposium*, 1996, pp. 137-144.
- [17] J. Klaue, "EvalVid - A Video Quality Evaluation Tool-set", in <http://www2.tkn.tu-berlin.de/research/evalvid/fw.html#bin>, Telecommunication and Network Group (TKN), Faculty of EE and CS, Berlin, 2003.
- [18] J. Klaue, "EvalVid - A Video Quality Evaluation Tool set, in <http://www2.tkn.tu-berlin.de/research/evalvid/cif.html>, Telecommunication and Network Group (TKN), Faculty of EE and CS, Berlin, 2003.

- [19] A. P. Patil, N. Sambaturu, and K. Chunhaviriyakul, ., "Convergence time evaluation of algorithms in MANETs", in *arXiv preprint arXiv:0910.1475*, 2009.

AUTHORS PROFILE

Sabrina Nefti is currently reading for Masters in Network Architecture at the University of HAL, Batna, Algeria. Certified Information Systems Auditor (UK, 2010). Previous studies and professional experience: MEng in Computer Engineering with First Class Honors from the University of Southampton, United Kingdom of Great Britain (2007), IT Senior Consultant at Ernst & Young (London, UK, 2007-2014), Assistant Engineer at Advanced Risc Machines Limited (Cambridge, UK, 2004-2005); Best Performance Award at MEng degree at Southampton University (UK, 2007), The Coulton Medal for Achievement in Measuring and Control (UK, 2006), Four Zepler Awards for Best Performance at Southampton University, UK (2002, 2003, 2006, 2007). Research interests: 5th generation communication protocols, mobile learning, Internet of Things, pervasive healthcare.

Maamar Sedrati received his engineering degree in 1985 from UMC Constantine, Algeria and he obtained the Ph.D. degree in 2011 from UHL Batna University, Algeria. He is currently serving as an assistant professor and member of LaSTIC laboratory at the Computer Science Department, University of Batna 2, Algeria. His research interests include computer networks, Internet technologies and mobile computing, security, quality of service for Multimedia applications in wireless and mobile networks and several aspects of the Internet of Things (IoT). He is a member of technical program committees in many national and international conferences such as ICACIS, CN2TI, IPAC and ICCSA.

Automatically Determining the Location and Length of Coronary Artery Thrombosis Using Coronary Angiography

Mahmoud Al-Ayyoub, Ala'a Oqaily and Mohammad I. Jarrah
Jordan University of Science and Technology
Irbid, Jordan

Emails: maalshbool@just.edu.jo, {alaa.oqaily, jarrahmohammad}@yahoo.com

Huda Karajeh
The University of Jordan
Amman, Jordan

Email: h.karajeh@ju.edu.jo

Abstract—Computer-aided diagnosis (CAD) systems have gained a lot of popularity in the past few decades due to their effectiveness and usefulness. A large number of such systems are proposed for a wide variety of abnormalities including those related to coronary artery disease. In this work, a CAD system is proposed for such a purpose. Specifically, the proposed system determines the location of thrombosis in x-ray coronary angiograms. The problem at hand is a challenging one as indicated by some researchers. In fact, no prior work has attempted to address this problem to the best of our knowledge. The proposed system consists of four stages: image preprocessing (which involves noise removal), vessel enhancement, segmentation (which is followed by morphological operations) and localization of thrombosis (which involves skeletonization and pruning before localization). The proposed system is tested on a rather small dataset and the results are encouraging with a 90% accuracy.

Keywords—Heterogeneous wireless networks, Vertical handoff, Markov model, Artificial intelligence, Mobility management.

I. INTRODUCTION

Computer-aided diagnosis (CAD) systems are interdisciplinary systems whose main objective is to aid the physicians in the diagnosis process by interpreting different symptoms and tests such as blood tests, biopsies, medical images, etc. From their name, CAD systems are used to assist the physicians by offering them some guidance in interpreting test results or localizing the region of interest (ROI) in images leaving the final decisions to them. They even operate in an interactive manner incorporating the physician's input into the guidance/suggestions they offer. This makes CAD systems different from automated computer diagnosis systems, in which the diagnosis depends only on the system's output.

CAD systems are considered to be domain-specific. They are optimized for certain types of diseases, diagnosis methods, parts of the body, etc. They analyze different kinds of input such as symptoms, laboratory tests, medical images, etc. The most familiar kind of CAD systems is the one that analyzes the medical images. They are considered challenging because they combine the elements of artificial intelligence and digital images processing. However, there is so much effort put into such systems to improve their efficiency and accuracy and to integrate them into the software of the medical imaging machines that they have spread widely and quickly in the radiology mainstream to provide quick and accurate diagnosis.

For example it has already become a part of the routine clinical work for the detection of breast cancer with mammograms at many hospitals in the United States. In general, CAD systems are beginning to be applied widely in detecting different types of abnormalities in medical images obtained by different modalities such as magnetic resonance imaging (MRI), computed tomography (CT), x-ray images, etc. [15], [16].

CAD systems are concerned with the computerized extraction of quantitative measurements from medical images. They usually consist of several steps including image preprocessing, feature extraction and analysis and classification based on the extracted features via the use of machine learning classifiers such as decision trees and Artificial Neural Network (ANN) [35]. Following such framework in this work, we design a CAD system that processes x-ray coronary angiograms to detect the location of thrombosis in coronary artery vessels.

Coronary artery disease is the most common cause of sudden death and one of the world's most important causes of early mortality. It is the most common reason for death of men and women over 20 years of age. Coronary artery disease diagnostic procedure is usually performed in a sequential manner ending with imaging of the coronary vessels that is performing coronary angiography test. This test is the most accurate test amongst all and is considered as the reference method to confirm the existence of coronary artery disease. In our work, we examine this test as a reference method to diagnose the coronary artery disease, in particular the detection of thrombosis in x-ray angiograms.

Thrombosis is the common cause of heart attack or myocardial infarction (MI) which if not treated properly can cause sudden and severe death. The heart is surrounded by three major arteries that supply it with blood and oxygen. MI is the case when one of these coronary arteries is obstructed by blood clot (thrombosis) leading to insufficient supply of blood and oxygen to the heart muscle which consequently results in tissue death (form a lesion called infarction). We concentrate on thrombosis that forms in the coronary arteries, causing MI or heart attack. These types of thrombosis are detected/visualized invasively by x-ray coronary angiography. The diagnosis process is considered to be very challenging as it requires highly experienced physicians. So any improvement in the diagnostic procedure is highly appreciated and required.

This research introduces a system for detecting the location of thrombosis so it can be helpful for improving the diagnostic process.

The diagnosis of thrombosis is a very challenging task as it needs lots of visual acuity and requires highly experienced cardiologists to confirm its existence and differentiate it from others coronary lesions. This is especially important due to the rough nature of x-ray images and the amount of noise and overlapping organs that may appear in the images and can interrupt the visualization of ROI in addition to other factors that may affect the quality of the images. Therefore, the existence of any diagnostic improvements that are efficient and accurate is highly appreciated. In this work, we design a CAD system for detecting thrombosis in coronary x-ray angiograms. To the best of our knowledge, this problem, which is the localization of coronary artery thrombosis or even the detection of thrombosis, has not been discussed in any previous work.

The proposed CAD system has many advantages. First, it offers the physicians a safe, reliable, efficient and accurate way requiring minimal effort to aid them in the diagnosis process. This will allow them to avoid human errors due to physiological factors, emotional problems, tiredness, stress, overworking, distractions, etc. Another advantage of the system is its potential benefits as a teaching/training tool for senior medical students and junior residents. Finally, a third advantage is gained by linking the proposed system with medical images storage and retrieval system. By doing so, researches would spend minutes collecting cases for their studies even if the available cases are in order of millions which would require months if the proposed system is not used.

The rest of this paper is organized as follows. The following section presents a review of the previous attempts to address related problems in the literature. Section III illustrates the methodology followed in this research and the steps undertaken by the proposed system in detailed manner. Section IV discusses the experiments and shows the resulting images from each stage in the proposed system. Finally, the conclusion and the future works are presented in Section V.

II. RELATED WORKS

This sections starts by giving a brief overview of the literature on CAD systems before discussing the works that are more related to the problem at hand. According to [15], using computers to analyze medical images dates back to the 1960s. Since then, CAD systems have received continuous attention from different researchers in academia and the industry. CAD systems vary greatly in the tasks they target and the information they use. While certain organs have received a lot of attention (such as breast [30], [13], [32], [6], [20], [18], chest [22], [41], [38], and colon [9], [40]), other organs (such as brain [5], [3], liver [25], [11], and skeletal and vascular systems [7], [4], [31]) are less studied. Another variation in CAD systems is their applicability. While some CAD systems have been strongly tied with the industry, other CAD systems are still being honed to be more practical by either enhancing its accuracy or performance using better algorithms or special hardware [12], [21], [37], [34], [43], [1].

The following paragraphs briefly discuss the previous work related to the problem at hand starting. We start with some of the integral image processing tasks such as vessel extraction and segmentation.

Vessels extraction from coronary angiography images is highly desirable in many applications like measuring the vessel width to detect the existence of stenosis or abnormal branching or tortuosity, etc. [8], [14]. Most of the works on coronary angiograms have focused on two main categories: the preprocessing (vessel enhancement) and segmentation (vessel segmentation) of coronary angiograms. According to [14], the literature review on coronary angiography image enhancement is very limited and it is focused on image enhancement for the sake of improvement of subsequent segmentation rather than improving the quality of visualization in clinics. However, for image segmentation, there have been many methods that visualize the vessels. These methods can be classified into: model based tracking and propagation, pattern recognition, neural network, fuzzy and artificial intelligence-based methods. In [14], the authors covered both the early and recent works related to vessel segmentation algorithms and techniques. They presented the algorithms that are specifically used for x-ray angiography and compared between them according to different aspects such as the support of several scales, the need to interact with the user, the ability to handle small vessels and junctions, etc.

Kirbas et al. [23] provide a survey of vessel extraction (segmentation) techniques and algorithms. The goal of this paper is to introduce early and recent vessel extraction techniques and algorithms for the practitioners. The paper classifies the existing research on vessel extraction techniques and algorithms by grouping researches that uses similar approaches in the task of vessel extraction into the same category, and creates separate categories for methods that are used extensively. Accordingly this paper introduces six main categories for vessel segmentation techniques and algorithms, including: (1) pattern recognition techniques, (2) model based approaches, (3) tracking-based approaches, (4) artificial intelligence-based approaches, (5) neural network based approaches and (6) miscellaneous tube-like object detection approaches. Some of these categories are further divided into subcategories based on their intensive use in the literature. The paper tries to introduce each segmentation category, gives a brief summarization of the papers that fall into this category and refer the readers to references for additional information. At the end of each section the paper presents a comparison between the methods introduced in this section through a table. The comparison includes input image type (such as XRA, MRA, CT, etc.), use of a prior knowledge, dimensionality, user interaction requirement, whether the method implies multi-scale techniques, result type such as centerline, vessel edges, and junctions, and whether the method segments the whole vessel tree or not. Other works on vessel segmentation and analysis include [26], [10], [19], [42], [28], [24], [36].

While most of the work focused on vessel extraction and image enhancement, Syeda-Mahmood et al. [39] focused more on feature extraction from coronary angiograms to support clinical decisions. It quickly assesses the disease in the current patient by exploiting the extracted features from images to end similar coronary angiograms. Specifically it ends the sim-

ilar coronary angiograms by extracting clinically meaningful features such as number of significant junction, thickness of arteries, number of trifurcations tortuosity, lengths of artery segments and lumen variations. Then, it uses a supervised learning method called Relevant Component Analysis (RCA) as it is useful for feature comparison in classification phase.

For more information about the vessel segmentation and feature extraction algorithms on 3D imaging modalities, one can refer to Lesage et al. [27], which reviewed the state of the art literature of vessel segmentation on computed tomography angiography (CTA) and magnetic resonance angiography (MRA) 3D imaging modalities. This review analyzed the literature according to three high level axes: appearance and geometric models, image features and extraction schemes. It did not provide detailed discussion; instead, it only provided high level discussions. Also, it did not provide a comparison between the mentioned works. In order to get a highly detailed categorization of existing works on vessel segmentation algorithms and techniques, the interested reader is referred to [23], which provided a comparison based study of the existing works.

To the best of our knowledge there has been no work in the literature that deals with thrombosis in coronary angiograms. According to our search in the Internet as well as through our reading of the surveys in the area of coronary angiograms processing field that include a lot of the previous and current works in this field, there have been no evidence of any research that handles thrombosis. Moreover, Kirbas et al. [23] mentioned that thrombosis segmentation is a challenging task without providing any references to any work in this topic. This supports our beliefs that there is no work that handle thrombosis detection in coronary angiograms.

III. PROPOSED APPROACH

In this paper, a system is proposed that takes an x-ray angiogram and determines the location of the ROI (thrombosis) in conspicuous way to the observer. It consists of the following four main stages as depicted in Figure 1.

- 1) The image preprocessing stage which starts with converting the image from RGB to gray scale as it is a prerequisite for most of the algorithms in the following stages. After that, standard filters are employed to remove the Gaussian and salt and pepper noise.
- 2) Vessel enhancement stage, which is essential for the purposes of vessel segmentation or image feature extraction. The Frangi filter [17], a widely used filter customized for such purposes, is used.
- 3) Image segmentation, in which Otsu's thresholding algorithm is used followed by morphological operations to enhance the segmentation and remove undesirable objects. Such an approach has proven its effectiveness in previous CAD systems based on gray scale images [6], [3]. Although the common purpose of segmentation is to extract features or ROI within the images and suppress everything else, in this work, we exploit the absence of ROI to localize thrombosis.
- 4) Finally, localization of thrombosis stage in which we use skeletonization algorithms to compute the

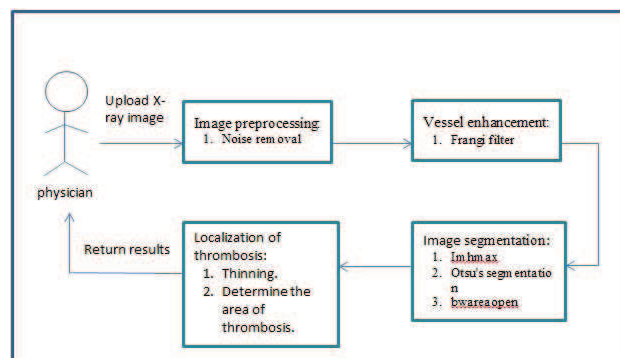


Fig. 1: Stages of the proposed system.

centerline/medial axis of the segmented images. The skeleton is then used to find the location of thrombosis. Finally, a circle is drawn on the original image indicating the location of thrombosis to provide visual aid to the physician.

Detailed discussions of these stages are given in the following subsections. Figure 2 shows six sample x-ray images from our dataset. Each step of the proposed system will be shown on these sample images.

A. Image Preprocessing

Image preprocessing refers to the process of enhancing images contrast, boundaries, noise reduction, sharpening of image features, filtering and other processes that aid the process of image display and analyses. In this work, we start by converting the image from RGB to gray scale, which means that the image will have pixel intensities in the range between 0 and 255. Then, noise removal tools are applied to prepare the images for the following steps.

Noise is the result of error in image acquisition affecting the true intensities of the real scene. Medical images in general, and specifically x-ray images, contain some noise [2]. The noise in such images gives them a mottled, grainy or snowy appearance, which reduces image quality and affects its usability. The effect of noise is most significant on relatively small objects with low contrast. In our work we will handle two type of noise, Gaussian noise and salt and pepper noise.

Gaussian Noise. Gaussian noise is a statistical noise where its probability density function (called Gaussian distribution) is equal to probability density function of the normal distribution. In order to solve the Gaussian noise we use the Wiener filter. It is a type of linear filters, but it produces results better than linear filters because it is an adaptive filter that preserves edges and other high frequency parts of the image. The Wiener filter is commonly used for noise reduction on an image or signal and it is known to perform well when the noise is Gaussian.

Salt and Pepper Noise. The salt and pepper or “speckle” noise is represented as black and white pixels randomly set in the image giving it the “salt and pepper” appearance. In order to remove the salt and pepper noise, the median filter is used. It is a nonlinear digital filtering technique that is widely used in

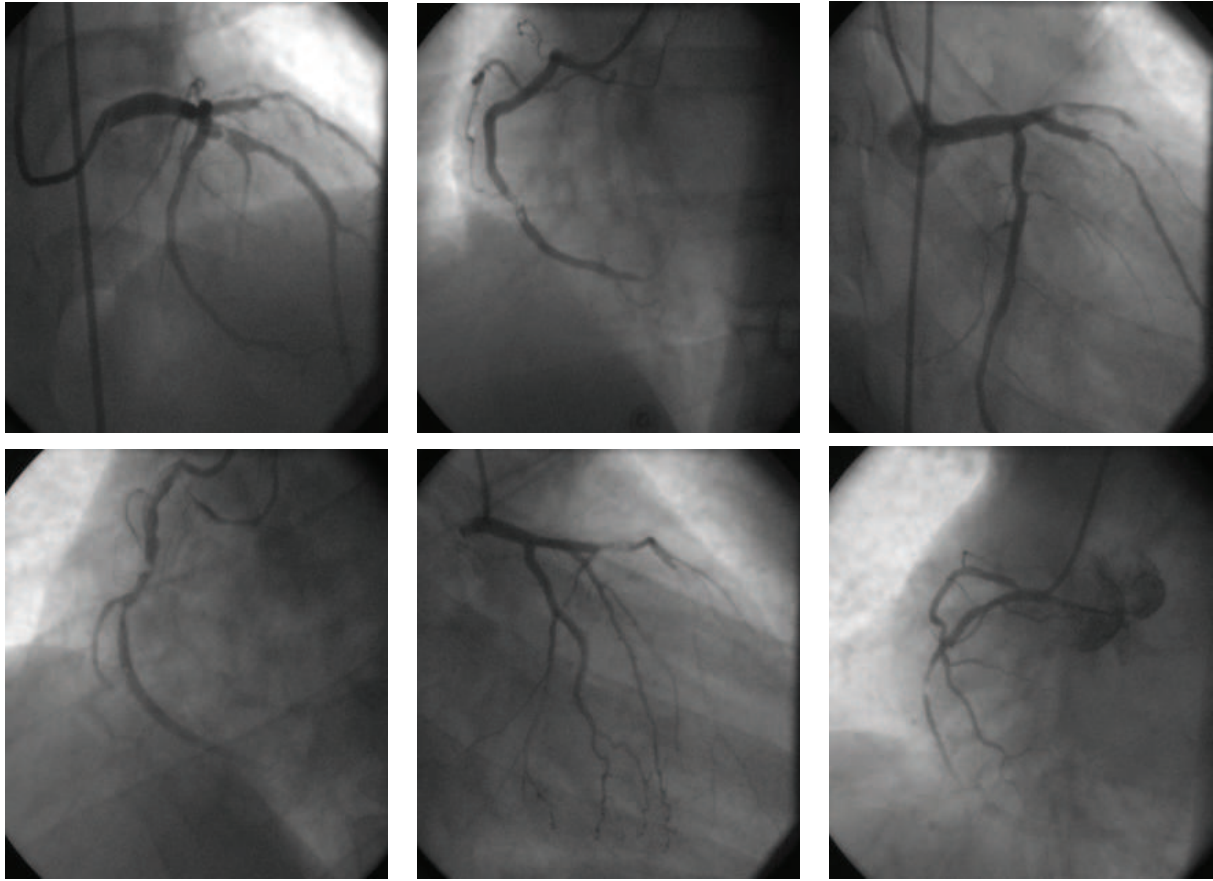


Fig. 2: Sample x-ray angiograms.

medical image processing because it is able to remove noise while preserving edges.

B. Vessel Enhancement

The task of vessel enhancement and extraction from x-ray angiograms is a challenging task due to many issues such as weak contrast between the coronary vessels and the background, strong overlapping shadows of the body organs and the bones, and easily deformable shape of the vessel tree. Moreover, the nature of the x-ray images can lead to varying image quality because of the varying contrast in the images due to factors like dye dose, patient weight, camera adjustments, etc., which could affect the quality of the images leading to more challenges in analyzing them.

Due to its importance, there have been several attempts to propose effective vessel enhancement and extraction algorithms. See [23], [27], [14] for surveys on most of these works. After deep search and thorough experimentation, one algorithm emerged as the top choice for this work based on its good performance and accurate results when applied to our dataset. This algorithm is the Frangi filter proposed by Frangi et al. [17]. It is used as a vessel enhancement algorithm with the ultimate goal of vessel segmentation. It supports both 2D images such as x-ray images as well as 3D images such as CTA. Due to its complexity, the Frangi filter will not be

explained in details in this paper. Only a high level description is provided in the following paragraph.

Similar to its predecessors (the works of Lorenz et al. [29] and Sato et al. [33]), the Frangi filter uses the eigenvalues of the Hessian matrix \mathcal{H} to derive structural information. One of the new things about it is that it uses all the eigenvalues ($|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$) of the Hessian to compute the likeliness of vessel existence. \mathcal{H} is computed using Gaussian derivatives of the image. The following equation gives the Gaussian second order derivative in point \vec{x}_o of image L at scale s .

$$\frac{\partial}{\partial x} L(\vec{x}, s) = s^\gamma L(\vec{x}) \times \frac{\partial}{\partial x} G(\vec{x}, s),$$

where where $G(\vec{x}, s)$, the D-dimensional Gaussian, is defined as follows.

$$G(\vec{x}, s) = \frac{1}{\sqrt{2\pi s^2}^D} e^{-\frac{\|\vec{x}\|^2}{2s^2}} \quad (1)$$

with γ being a normalization parameter. The Frangi filter uses the following vesselness function that works in a multi-scale framework in which the maximum and minimum scales at which vessels are expected to be found are denoted by s_{\min} and s_{\max} .

$$\mathcal{V}_o(\gamma) = \max_{s_{\min} \leq s \leq s_{\max}} \mathcal{V}_o(s, \gamma).$$

Moreover, the Frangi filter makes use of the following two

measures based on the second order ellipsoid

$$\mathcal{R}_B = \frac{\text{Volume}/(4\pi/3)}{(\text{Longest cross section area}/\pi)^{3/2}} = \frac{|\lambda_1|}{\sqrt{|\lambda_2\lambda_3|}} \quad (2)$$

$$\mathcal{R}_A = \frac{\text{Longest cross section area}/\pi}{(\text{Largest axis semi-length})^2} = \frac{|\lambda_2|}{|\lambda_3|} \quad (3)$$

where \mathcal{R}_B quantifies the deviation from a blob-like structure and \mathcal{R}_A help in distinguishing between plate-like and line-like structures. To deal with background noise, the Frangi filter uses the Frobenius matrix norm of the Hessian defined as follows.

$$\mathcal{S} = \|\mathcal{H}\|_F = \sqrt{\sum_{j \leq D} \lambda_j^2}, \quad (4)$$

where D is the dimension of the image. Finally, combining the measures defined in Equations 3-4 gives the following vesselness measure.

$$\mathcal{V}_o(s) = \begin{cases} 0 & \text{if } \lambda_2 > 0 \\ & \text{or } \lambda_3 > 0, \\ \left(1 - e^{-\frac{\mathcal{R}_A^2}{2\alpha^2}}\right) e^{-\frac{\mathcal{R}_B^2}{2\beta^2}} \left(1 - e^{-\frac{\mathcal{S}^2}{2c^2}}\right) & \text{otherwise} \end{cases}$$

where α , β and c are user-defined parameters to control the sensitivity of the line filter to the measures \mathcal{R}_A , \mathcal{R}_B and \mathcal{S} . Figure 3 shows the results of applying the Frangi filter on the images of Figure 2 after applying the noise removal filters.

C. Image Segmentation

Image segmentation is the process of partitioning an image into distinct regions containing pixels having similar attributes. It is an integral step in many CAD systems based on image analysis. In this work, Otsu's thresholding algorithm is used followed by morphological operations to enhance the segmentation and remove undesirable objects. Such an approach has proven its effectiveness in previous CAD systems [6], [3]. Otsu's method is a clustering-based image thresholding technique. In its most basic form, Otsu's method assumes that the image to be segmented is composed of two classes, the foreground and the background, and computes the threshold that minimizes the intra-class variance. Thus, it performs best when the image histogram is bi-modal. The image is segmented based on this global threshold. Figure 4 shows the results of applying the Otsu's segmentation algorithm on the images of Figure 3.

For the segmentation to give useful results for the problem at hand, morphological operations are used. As discussed in [6], morphological operations process geometrical structures and analyze them based on topology, set theory, lattice theory and random functions. They are widely used for different images processing tasks.

Erosion and dilation are two basic morphological operations. They are defined by Equations 5 and 6, respectively.

$$A \ominus B = \{z | (B)_z \subset A\} \quad (5)$$

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (6)$$

where A and B are the image and the structure element, respectively. Afterwards, objects that are not related to the vessel tree are removed by computing the connected components in the image (which requires converting it into a binary image) and discarding small components.

D. Localization of Thrombosis

The resulting image of the previous steps is a binary one representing the coronary artery vessels. The affected area (the coronary artery thrombosis) is shown as a disconnection in the artery vessel. Hence, we focus on finding such disconnections. To simplify the task at hand, the vessels are reduced to simple "thin" lines representing their medial access.

Extracting the Vessels Centerlines. In order to extract the vessel centerline/medial axis, skeletonization (skeleton extraction) algorithms are used. Skeletons and medial axes are used interchangeably in the literature; however, some researchers discuss the subtle differences between them. Here, we make no such distinction. Even though there are different variants of skeletonization techniques, the experiments conducted on our collected dataset revealed that both morphological skeletonization algorithm (based on morphological opening) and the thinning algorithm (based on the hit-and-miss transform) perform well. However, we decide to use the thinning algorithm as its produced skeletons are more straightened and tidier than those produced by the morphological openings leading to better accuracy by the localization algorithm. Figure 5 shows the results of applying the skeletonization algorithm on the images of Figure 4 after applying the morphological operations.

Note that after applying the thinning algorithm, undesirable short spurs (branches of a skeleton line which are not key to the overall shape of the line) appear. Such spurs are produced by small irregularities in the boundary of the vessels. They affect the accuracy of our localization algorithm and they are removed using a pruning algorithm.

Determining Thrombosis Location. In this step, we determine the location of thrombosis by exploiting the idea that it splits the artery vessel into two separate segments. This is computed through finding the end points of the skeleton lines representing the artery vessels, and finding the closest pair of endpoints that do not belong to the same skeleton line. Once again, we will binary morphological operations to extract the end points of the skeleton lines. After finding these end points we can easily sketch a circle indicating the location of thrombosis on the original image. Figure 6 shows the results of applying the localization algorithm on the images of Figure 5.

IV. EXPERIMENT AND RESULTS

An evaluation of our proposed algorithm is presented in this section. The dataset used in this research is collected from the Cath Lab at King Abdullah University Hospital (KAUH) and annotated by human experts from the same lab. It consists of 20 x-ray images acquired from 20 patients. They contain different views that include the three major coronary arteries: left anterior descending (LAD), left circumflex (LCX) and right coronary artery (RCA). The process of collecting data is hard during to the rarity of thrombosis cases at KAUH and the inefficient image storage and retrieval system; a problem that will hopefully be solved by adopting our proposed system. For each case in the dataset, an experienced physician manually indicates the location of thrombosis. The experiments show that our proposed system is successful in 90% of the times despite the varying quality of the dataset images. Figure 6 shows the output of our system for the six cases of Figure 2

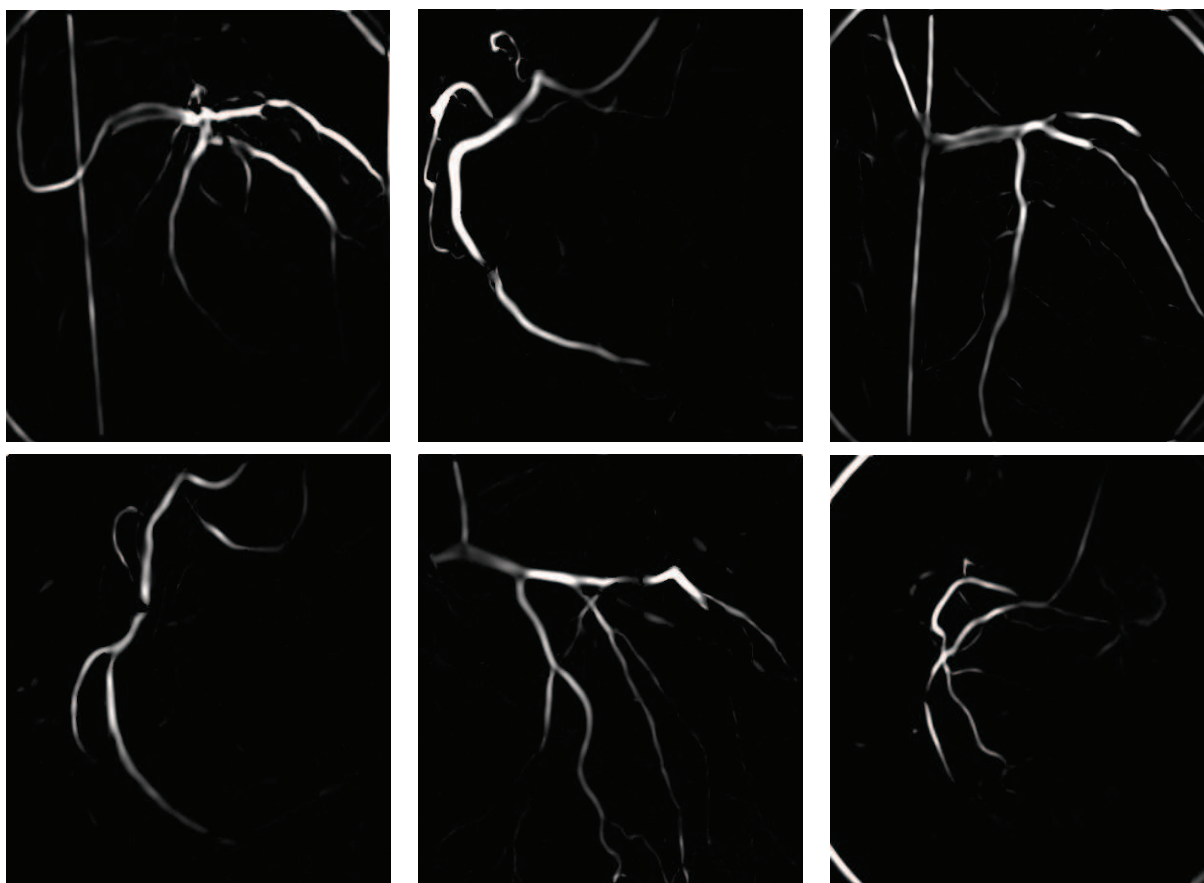


Fig. 3: The results of applying the Frangi filter on the images of Figure 2 after applying the noise removal filters.

and it can be seen that the system managed to automatically localize thrombosis in a very accurate manner. The only failures the proposed system face is for two difficult cases in which thrombosis does not appear clearly even to the human expert.

According to our human experts, the length of the thrombosis can play a significant impact on the treatment path taken by the cardiologists. So, As an added feature to our proposed system, the length of the thrombosis is computed. Figure 7 2 shows sample images from the dataset proving that our system can be very accurate in determining the length of thrombosis where the different between the length of the thrombosis region as computed by our system and the length of the thrombosis region computed based on the human experts' annotations does not reach 10%. However, there are cases for which the proposed system performs less perfectly such as the one depicted in Figure 8. Focusing on improving the system's accuracies for such cases is one of the future directions of this work.

V. CONCLUSION AND FUTURE WORK

In this work, a CAD system is proposed for determining the location of thrombosis in x-ray coronary angiograms. The problem at hand is a challenging one as indicated by some researchers. In fact, no prior work has attempted to address this problem to the best of our knowledge. The proposed system

consists of four stages: image preprocessing (which involves noise removal), vessel enhancement, segmentation (which is followed by morphological operations) and localization of thrombosis (which involves skeletonization and pruning before localization). The proposed system is tested on an in-house dataset and the results are impressive with a 90% accuracy. Moreover, the assist the cardiologists determine the best treatment option, the proposed system accurately computes the length of the thrombosis region.

This work is far from its final stages. After designing a system for automatically locating and measuring the thrombosis region and establishing its accuracy, the next step is to integrate it into the software of the medical imaging machine as well as the medical image storage and retrieval systems which would allow physicians and researchers to take full advantage of it.

VI. ACKNOWLEDGMENTS

The authors would like to thank the Deanship of Research at the Jordan University of Science and Technology for supporting this work (Project No. 20140048).

REFERENCES

- [1] Mahmoud Al-Ayyoub, Ansam M Abu-Dalo, Yaser Jararweh, Moath Jarrah, and Mohammad Al Sad. A gpu-based implementations of the fuzzy c-means algorithms for medical image segmentation. *The Journal of Supercomputing*, 71(8):3149–3162, 2015.

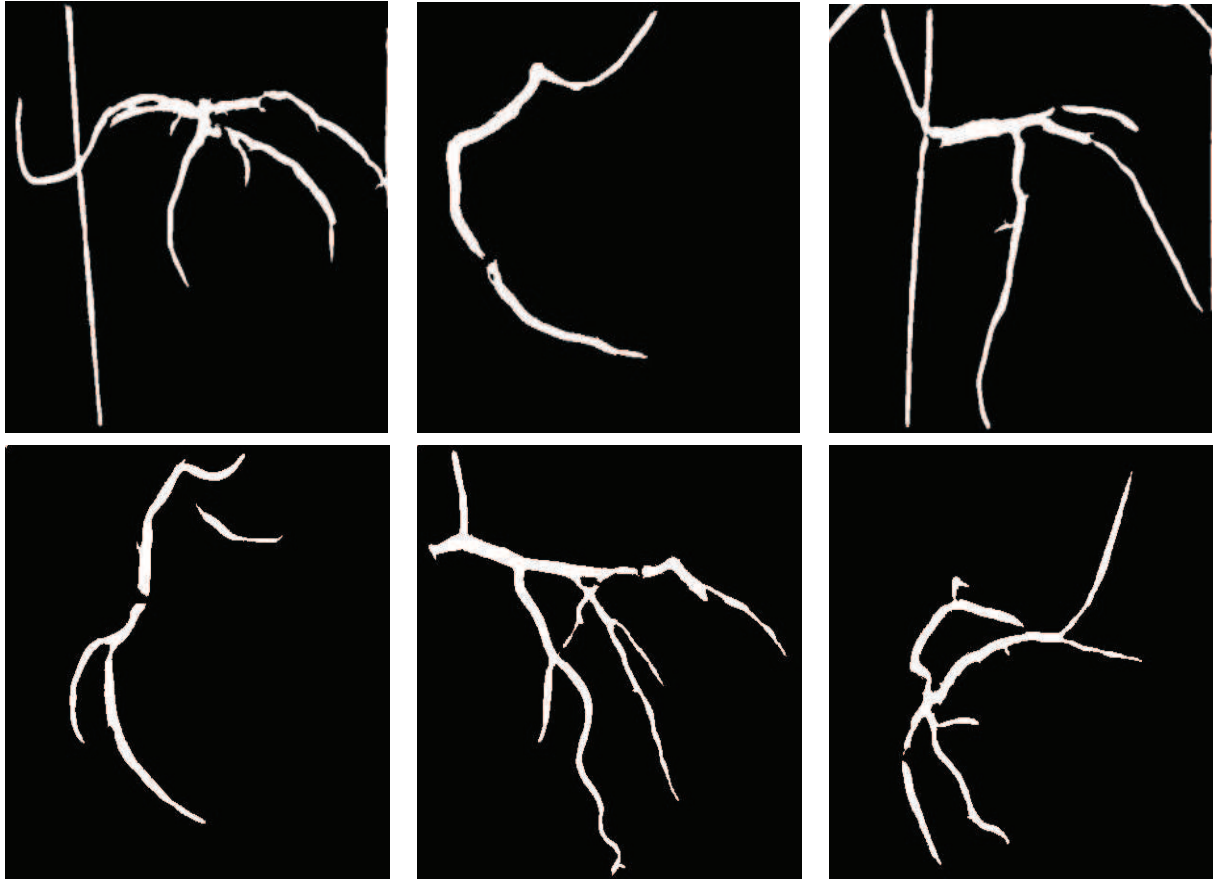


Fig. 4: The results of applying the Otsu's segmentation algorithm on the images of Figure 3.

- [2] Mahmoud Al-Ayyoub and Duha Al-Zghool. Determining the type of long bone fractures in x-ray images. *WSEAS Transactions on Information Science and Applications*, 10(8):261–270, 2013.
- [3] Mahmoud Al-Ayyoub, Duaa Alawad, Khaldun Al-Darabsah, and Inad Aljarrah. Automatic detection and classification of brain hemorrhages. *WSEAS Transactions on Computers*, 12(10):395–405, 2013.
- [4] Mahmoud Al-Ayyoub, Ismail Hmeidi, and Haya Rababah. Detecting hand bone fractures in x-ray images. *Journal of Multimedia Processing and Technologies (JMPT)*, 4(3):155–168, 2013.
- [5] Mahmoud Al-Ayyoub, Ghaith Husari, Omar Darwish, and Ahmad Alabed-alaziz. Machine learning approach for brain tumor detection. In *Proceedings of the 3rd International Conference on Information and Communication Systems (ICICS 2012)*, 2012.
- [6] Khaldun Al-Darabsah and Mahmoud Al-Ayyoub. Breast cancer diagnosis using machine learning based on statistical and texture features extraction. In *Proceedings of the 4th International Conference on Information and Communication Systems (ICICS 2013)*, 2013.
- [7] Khaled Alawneh, Mays Al-dwiekat, Mohammad Alsmirat, and Mahmoud Al-Ayyoub. Computer-aided diagnosis of lumbar disc herniation. In *Information and Communication Systems (ICICS), 2015 6th International Conference on*, pages 286–291. IEEE, 2015.
- [8] Elisabeth Arnoldi, Mulugeta Gebregziabher, U Joseph Schoepf, Roman Goldenberg, Luis Ramos-Duran, Peter L Zwerner, Konstantin Nikolaou, Maximilian F Reiser, Philip Costello, and Christian Thilo. Automated computer-aided stenosis detection at coronary ct angiography: initial experience. *European radiology*, 20(5):1160–1167, 2010.
- [9] L Bogoni, P Cathier, M Dundar, A Jerebko, S Lakare, J Liang, S Periaswamy, ME Baker, and M Macari. Computer-aided detection (cad) for ct colonography: a tool to address a growing need. *The British Journal of Radiology*, 2014.
- [10] Terrence Chen, Gareth Funka-Lea, and Dorin Comaniciu. Robust and fast contrast inflow detection for 2d x-ray fluoroscopy. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 243–250. Springer, 2011.
- [11] Yen-Wei Chen, Jie Luo, Chunhua Dong, Xianhua Han, Tomoko Tateyama, Akira Furukawa, and Shuzo Kanasaki. Computer-aided diagnosis and quantification of cirrhotic livers based on morphological analysis and machine learning. *Computational and mathematical methods in medicine*, 2013, 2013.
- [12] Jae Young Choi, Dae Hoe Kim, Konstantinos N Plataniotis, and Yong Man Ro. Classifier ensemble generation and selection with multiple feature representations for classification applications in computer-aided detection and diagnosis on mammography. *Expert Systems with Applications*, 46:106–121, 2016.
- [13] Judy C Dean and Christina C Ilvento. Improved cancer detection using computer-aided detection with diagnostic and screening mammography: prospective study of 104 cancers. *American Journal of Roentgenology*, 187(1):20–28, 2006.
- [14] Maryam Taghizadeh Dehkordi, Saeed Sadri, and Alimohamad Doost-hoseini. A review of coronary vessel segmentation algorithms. *Journal of medical signals and sensors*, 1(1):49, 2011.
- [15] Kunio Doi. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, 31(4-5):198–211, 2007.
- [16] Kunio Doi. Current status and future potential of computer-aided diagnosis in medical imaging. *The British journal of radiology*, 2014.
- [17] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 1998*, pages 130–137. Springer, 1998.
- [18] Kavita Ganesan, URajendra Acharya, Chua Kuang Chua, Lim Choo Min, K Thomas Abraham, and Kung Bo Ng. Computer-aided breast

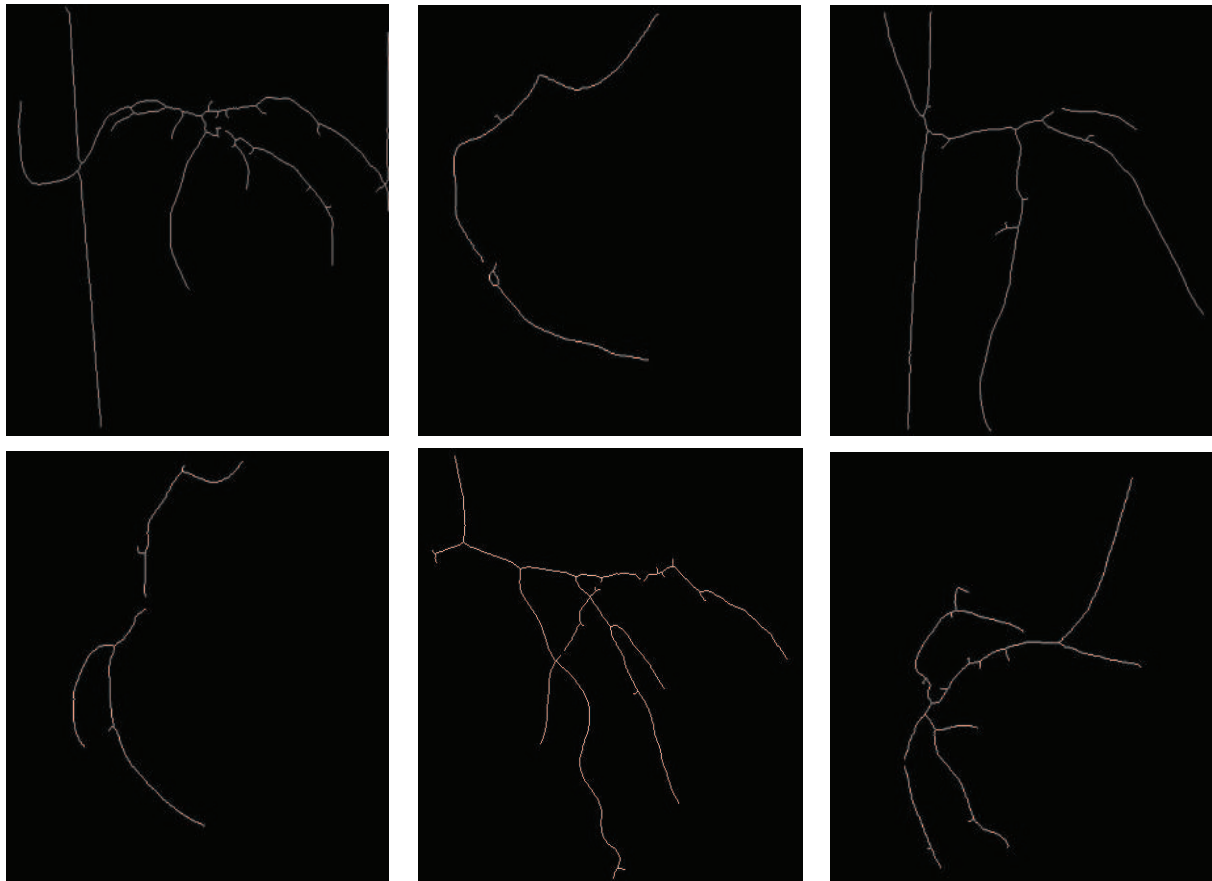


Fig. 5: The results of applying the skeletonization algorithm on the images of Figure 4 after applying the morphological operations.

cancer detection using mammograms: a review. *Biomedical Engineering, IEEE Reviews in*, 6:77–98, 2013.

[19] Vipin Gupta, Amit Kale, and Hari Sundar. A robust and accurate approach to automatic blood vessel detection and segmentation from angiography x-ray images using multistage random forests. In *SPIE Medical Imaging*, pages 83152F–83152F. International Society for Optics and Photonics, 2012.

[20] Afsaneh Jalalian, Syamsiah BT Mashohor, Hajjah Rozi Mahmud, M Iqbal B Saripan, Abdul Rahman B Ramli, and Babak Karasfi. Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. *Clinical imaging*, 37(3):420–426, 2013.

[21] Moath Jarrah, Muneera Al-Quraan, Yaser Jararweh, and Mahmoud Al-Ayyoub. Medgraph: a graph-based representation and computation to handle large sets of images. *Multimedia Tools and Applications*, pages 1–17, 2016.

[22] Yulei Jiang, Robert M Nishikawa, Robert A Schmidt, Charles E Metz, Maryellen L Giger, and Kunio Doi. Improving breast cancer diagnosis with computer-aided diagnosis. *Academic radiology*, 6(1):22–33, 1999.

[23] Cemil Kirbas and Francis Quek. A review of vessel extraction techniques and algorithms. *ACM Computing Surveys (CSUR)*, 36(2):81–121, 2004.

[24] Cemal Köse and Cevat İkibaş. Extraction of coronary vessel structures in low quality x-ray angiogram images. In *The Second International Conference Problems of Cybernetics and Informatics*, pages 240–243, Baku, Azerbaijan, September 2008.

[25] SS Kumar, RS Moni, and J Rajeesh. An automatic computer-aided diagnosis system for liver tumours on computed tomography images. *Computers & Electrical Engineering*, 39(5):1516–1526, 2013.

[26] Daniel SD Lara, Alexandre WC Faria, A de A Araujo, and David Menotti. A semi-automatic method for segmentation of the coronary artery tree from angiography. In *Computer Graphics and Image Processing (SIBGRAPI), 2009 XXII Brazilian Symposium on*, pages 194–201. IEEE, 2009.

[27] David Lesage, Elsa D Angelini, Isabelle Bloch, and Gareth Funka-Lea. A review of 3d vessel lumen segmentation techniques: Models, features and extraction schemes. *Medical image analysis*, 13(6):819–845, 2009.

[28] Chih-Yang Lin and Yu-Tai Ching. Extraction of coronary arterial tree using cine x-ray angiograms. *Biomedical Engineering: Applications, Basis and Communications*, 17(03):111–120, 2005.

[29] Cristian Lorenz, I-C Carlsen, Thorsten M Buzug, Carola Fassnacht, and Jürgen Weese. A multi-scale line filter with automatic scale selection based on the hessian matrix for medical image segmentation. In *Scale-Space Theory in Computer Vision*, pages 152–163. Springer, 1997.

[30] Marilyn J Morton, Dana H Whaley, Kathleen R Brandt, and Kimberly K Amrami. Screening mammograms: Interpretation with computer-aided detectionprospective evaluation 1. *Radiology*, 239(2):375–383, 2006.

[31] Ala’a Oqaily, Mohammad I Jarrah, Huda Karajeh, Mahmoud Al-Ayyoub, and Ismail Hmeidi. Localization of coronary artery thrombosis using coronary angiography. In *The Third international conference on informatics engineering and information science (ICIEIS2014)*, pages 310–316. The Society of Digital Information and Wireless Communication, 2014.

[32] Rangaraj M Rangayyan, Fabio J Ayres, and JE Leo Desautels. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, 344(3):312–348, 2007.

[33] Yoshinobu Sato, Shin Nakajima, Nobuyuki Shiraga, Hideki Atsumi, Shigeyuki Yoshida, Thomas Koller, Guido Gerig, and Ron Kikinis. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Medical image analysis*, 2(2):143–168, 1998.

[34] Mohammed A Shehab, Mahmoud Al-Ayyoub, and Yaser Jararweh. Improving fcm and t2fcm algorithms performance using gpus for

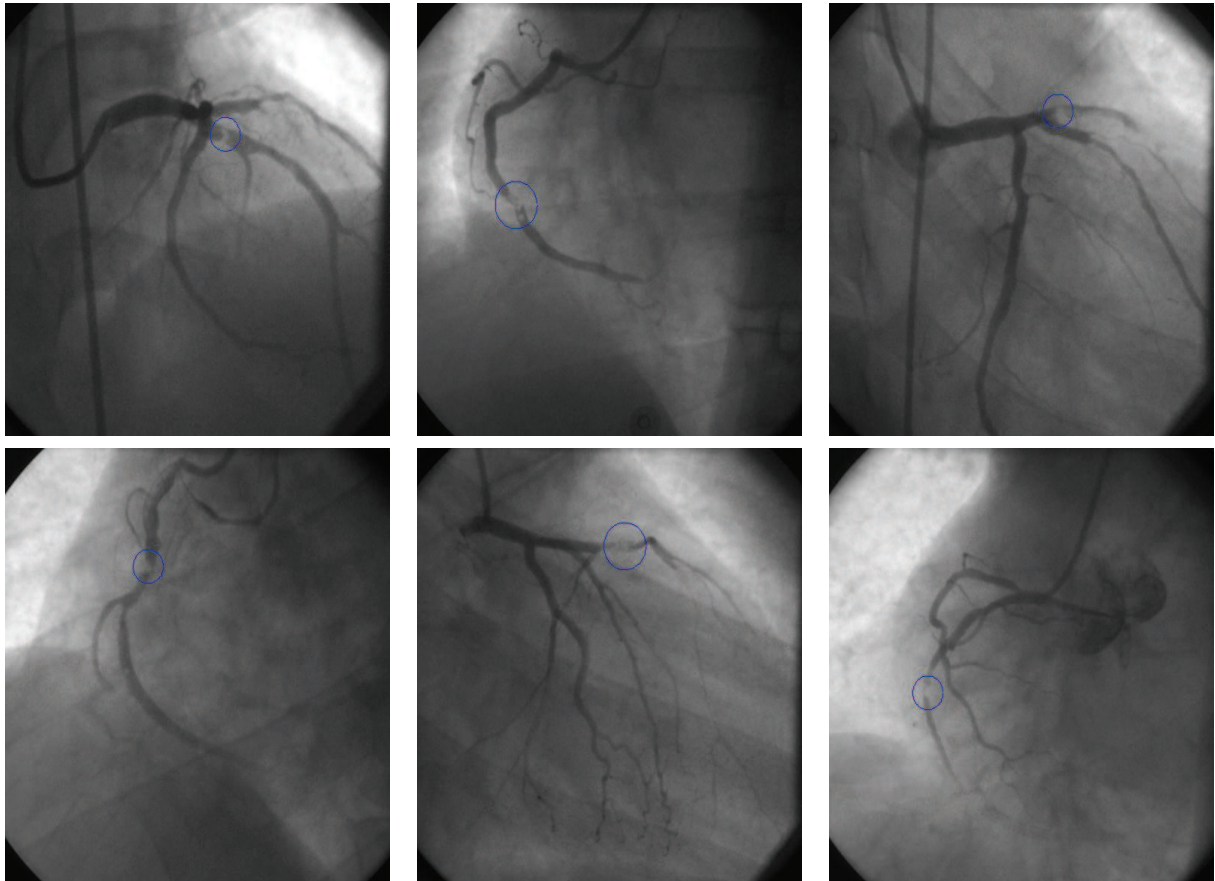


Fig. 6: The results of applying the localization algorithm on the images of Figure 5.

- medical images segmentation. In *Information and Communication Systems (ICICS), 2015 6th International Conference on*, pages 130–135. IEEE, 2015.
- [35] Junji Shiraishi, Qiang Li, Daniel Appelbaum, and Kunio Doi. Computer-aided diagnosis and artificial intelligence in clinical imaging. In *Seminars in nuclear medicine*, volume 41, pages 449–462. Elsevier, 2011.
- [36] Zhou Shoujun, Yang Jian, Wang Yongtian, and Chen Wufan. Automatic segmentation of coronary angiograms based on fuzzy inferring and probabilistic tracking. *Biomedical engineering online*, 9(1):40, 2010.
- [37] Konstantinos P Sidiropoulos, Spiros A Kostopoulos, Dimitris T Glotsos, Emmanouil I Athanasiadis, Nikos D Dimitropoulos, John T Stonham, and Dionisis A Cavouras. Multimodality gpu-based computer-assisted diagnosis of breast cancer using ultrasound and digital mammography images. *International journal of computer assisted radiology and surgery*, 8(4):547–560, 2013.
- [38] Tao Sun, Jingjing Wang, Xia Li, Pingxin Lv, Fen Liu, Yanxia Luo, Qi Gao, Huiping Zhu, and Xiuhua Guo. Comparative evaluation of support vector machines for computer aided diagnosis of lung cancer in ct based on a multi-dimensional data set. *Computer methods and programs in biomedicine*, 111(2):519–524, 2013.
- [39] Tanveer Syeda-Mahmood, Fei Wang, Ritwik Kumar, David Beymer, Y Zhang, Robert Lundstrom, and Edward McNulty. Finding similar 2d x-ray coronary angiograms. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 501–508. Springer, 2012.
- [40] SA Taylor, D Burling, M Roddie, L Honeyfield, J McQuillan, P Bassett, and S Halligan. Computer-aided detection for ct colonography: incremental benefit of observer training. *The British journal of radiology*, 2014.
- [41] Bram Van Ginneken, Bart M ter Haar Romeny, and Max A Viergever. Computer-aided diagnosis in chest radiography: a survey. *Medical Imaging, IEEE Transactions on*, 20(12):1228–1241, 2001.
- [42] Yuan Wang, Christine Toumoulin, HZ Shu, ZD Zhou, and Jean-Louis Coatrieux. Vessel extraction in coronary x-ray angiography. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 1584–1587. IEEE, 2005.
- [43] Terry S Yoo, Bradley C Lowekamp, Oleg Kuybeda, Kedar Narayan, Gabriel A Frank, Alberto Bartesaghi, Mario Borgnia, Sriram Subramaniam, Guillermo Sapiro, and Michael J Ackerman. Accelerating discovery in 3d microanalysis: Leveraging open source software and desktside high performance computing. *Microscopy and Microanalysis*, 20(S3):774–775, 2014.

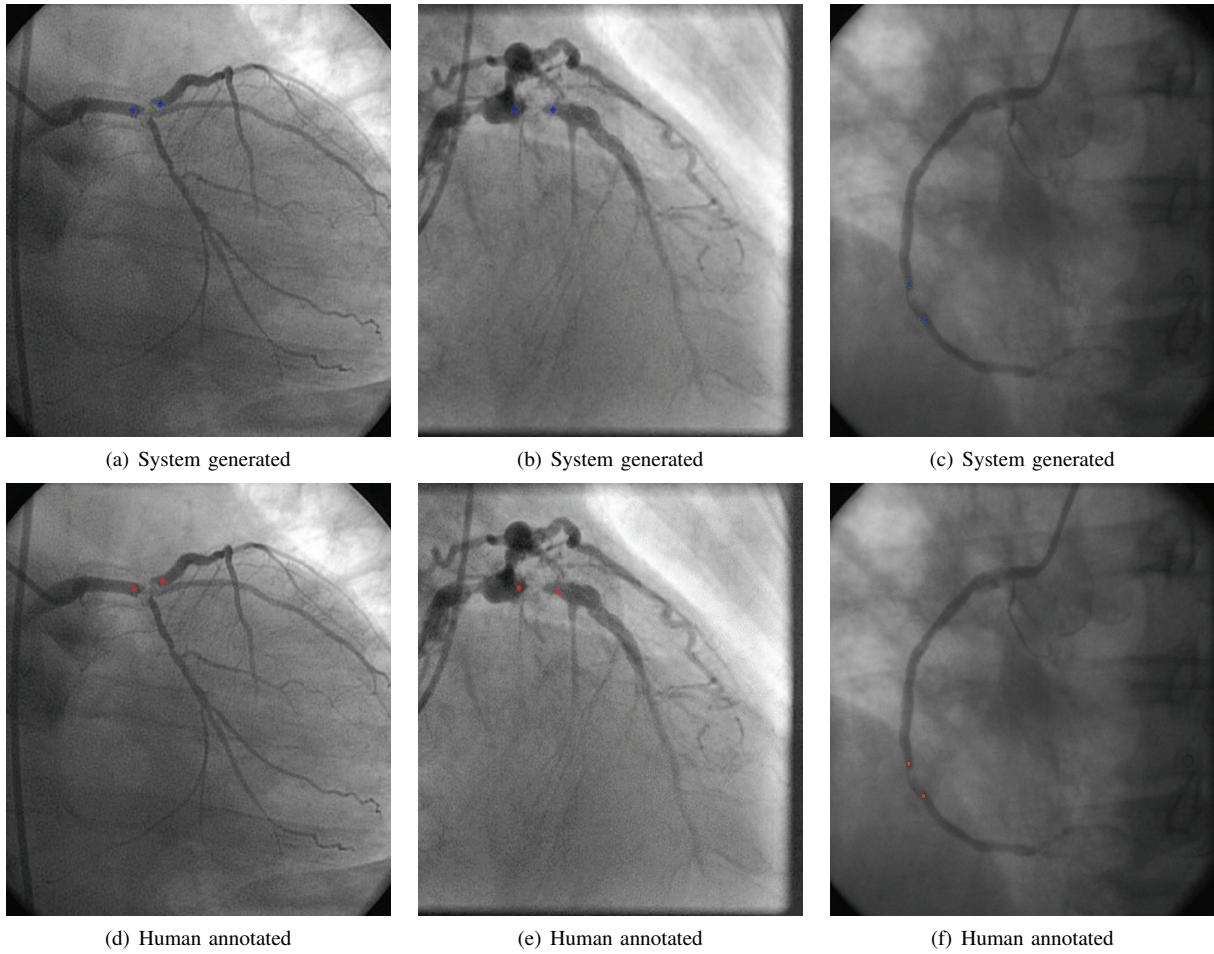


Fig. 7: Comparing the results of the proposed system (the blue dots in the top row of images) with human expert annotation (the red dots in the bottom row of images).

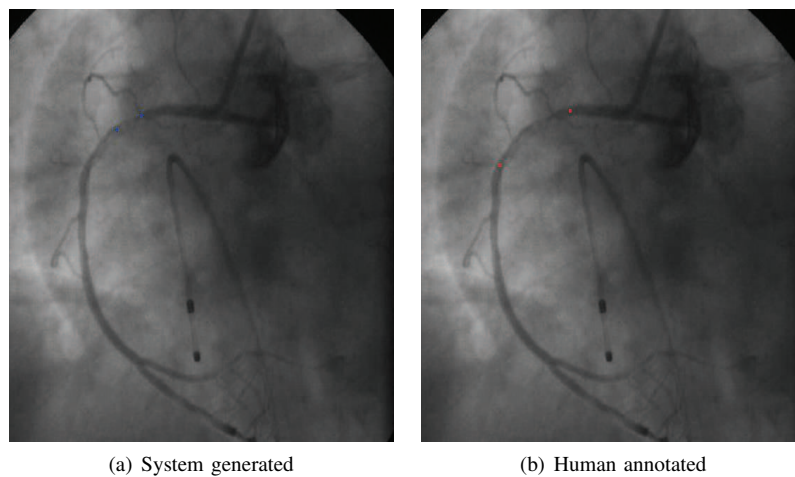


Fig. 8: A bad case for the proposed system.

Neutralizing vulnerabilities in Android: a process and an experience report

Carlos André Batista de Carvalho ^{#*1}, Rossana Maria de Castro Andrade ^{*2}, Márcio E. F. Maia ^{*3},
Davi Medeiros Albuquerque ^{*4}, Edgar Tarton Oliveira Pedrosa ^{*5}

[#] *Computer Science Department, Federal University of Piauí, Brazil*

¹ *candrebc@ufpi.edu.br*

^{*} *Group of Computer Networks, Software Engineering, and Systems, Federal University of Ceará, Brazil*

² *rossana@ufc.br*

³ *marcio@great.ufc.br*

⁴ *davialbuquerque@great.ufc.br*

⁵ *edgarpedrosa@great.ufc.br*

Abstract—Mobile devices became a natural target of security threats due their vast popularization. That problem is even more severe when considering Android platform, the market leader operating system, built to be open and extensible. Although Android provides security countermeasures to handle mobile threats, these defense measures are not sufficient and attacks can be performed in this platform, exploiting existing vulnerabilities. Then, this paper focuses on improving the security of the Android ecosystem with a contribution that is two-fold, as follows: i) a process to analyze and mitigate Android vulnerabilities, scrutinizing existing security breaches found in the literature and proposing mitigation actions to fix them; and ii) an experience report that describes four vulnerabilities and their corrections, being one of them a new detected and mitigated vulnerability.

I. INTRODUCTION

The vast popularization of mobile devices and the current trend to use them as both personal and business devices have turned smartphones and tablets into a natural target of security threats. That problem becomes specially relevant in the Android platform, where the number of device shipment has pushed past 84% of the market share in third quarter of 2015¹. Following the constant sales growth in Android-based mobile devices, the number of malwares targeting mobile devices rises as well [1].

In addition, the openness of the Android ecosystem produces a security concern. By design principle, Android is built to permit any developer and manufacturer to contribute to the platform. In that direction, Android facilitates the customization of OS versions, permits applications to be distributed from alternative app stores, and facilitates device rooting [2]. That flexibility comes with the price of introducing security breaches.

Android provides important security countermeasures to handle mobile threats. These measures include application isolation and the introduction of a permission system. Application isolation is present to permit application data to be secured from other applications, while the permission system exists to control the resources accessed by an application.

Although these security mechanisms in Android are robust and in constant update, many security vulnerabilities have been found [3]. In this context, researchers in the industry and academy have been investigating causes and impact of security vulnerabilities, aiming to improve the overall Android security [4], [5], [3], [6], [7].

In that direction, this paper focuses on improving the security of the Android ecosystem with a contribution that is two-fold, as follows: i) a process to analyze and mitigate Android vulnerabilities, scrutinizing existing security breaches found in the literature and proposing mitigation actions to fix them; and ii) an experience report that describes four vulnerabilities and their correction, being one of them a new detected and mitigated vulnerability.

The process presented in this paper is a result of a partnership between academy and industry, and is applied to analyze existing security breaches found in the literature. Although the analyzed vulnerabilities are known, some of them can still be exploited. The vulnerabilities analysis is shown in Attacks Trends Reports (ATRs), which presents the reproduction steps, discusses the vulnerability impact, and suggests mitigation actions, based on the root cause of each vulnerability.

Among the vulnerabilities listed in the ATRs, four are described in this paper. The first one is a non-cataloged vulnerability discovered by our team that permits malicious applications to access the cryptographic keys used in a Virtual Private Network (VPN) connection, allowing it to log and decipher all data transmitted over the active VPN. Additionally, two root vulnerabilities are described, along with a cross-signed certificate generating a looped certificate chain, a scenario not well-handled by Android. These vulnerabilities were fixed, when we implemented the mitigation actions proposed by us.

In this paper, the security characteristics of the Android platform are shown in Section II, and the related work in Section III. In Section IV, the process employed for vulnerability analysis is described and the ATRs are summarized. Due to space limitations, only the fixed vulnerabilities were detailed in Section V, including the non-cataloged vulnerability. Final considerations are presented in Section 5, with future work

¹<http://www.gartner.com/newsroom/id/3169417>

suggestions.

II. ANDROID SECURITY

Security is an important principle considered in the design of computers systems. On top of the security mechanisms inherited from Linux, the Android operating system reuses mechanisms from its technological environment (e.g. hardware, programming language and mobile carrier infrastructure) and incorporates new ones (e.g. application permissions) [8]. Additionally, the Android security mechanisms are constantly monitored and new breaches are found periodically.

Unknown security breaches are explored in an eventual attack, affecting security properties of the Android system. In this research, several vulnerabilities in the Android Software Stack are analyzed. Thus, to propose a mitigation action, the root cause of each vulnerability is identified. Hence, it is necessary to understand the security model employed in the Android Software Stack, presented in Figure 1.

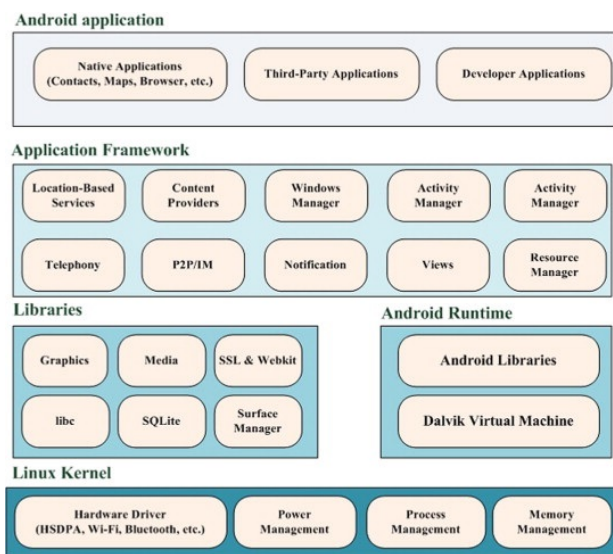


Fig. 1. Android Software Stack [9]

The *Application Layer* is the top layer of the Android software stack. Users interact with this layer to access device resources, such as to make phone calls and sending/receiving SMSs, or to use user installed applications (apps). The users should install applications using the Google Play Store² or device producer stores (e.g. LG SmartWorld³). Additionally, Android permits the installation of apps from other sources, such as non-official stores (e.g. I-Mobile⁴), Android Debug Bridge (ADB), Sideload or downloading of the app APK (Android Package).

The Android security model is based on application isolation, preventing sensitive information from being exposed to other applications. This isolation is implemented using device resource access managed by a permission system. An

Android developer must inform the application permissions and the user must authorize that access explicitly when the application is installed. With the permission system, device resources are protected and can be accessed only by system applications, either applications present in the Android Open Source Project (AOSP) or applications developed by the original equipment manufacturer (OEM). However, this control can be bypassed when the device is rooted, a process with which user applications obtain privileged control over restricted resources. For instance, user installed applications in rooted devices may capture Internet packages sent by other applications (see subsection V-A).

Android applications are executed in a sandbox, where Inter-Process Communication (IPC) mechanisms are required to allow application data to be accessed by other applications. Content Providers manage this communication that is available in the *Application Framework* layer. In this layer, there is an Application Programming Interface (API) providing a set of functions to be used in application development. Here, one important source of vulnerabilities is introduced by application developers, neglecting the security guidelines for application development [10]. Additionally, it is possible to find vulnerabilities in the Android Application Framework, such as the Cross-Signed Certificate, described in subsection V-D.

The next layer contains the *Android Native Libraries* and the *Android Runtime*. The native library is written in C/C++ and compiled, specifically, for each device. Recently, a vulnerability in the OpenSSL library, known as Freak SSL⁵, was discovered. This vulnerability allows the decryption and modification of SSL messages.

The Dalvik Virtual Machine and Core Libraries are in the Android Runtime. The Dalvik VM is optimized to devices with power and memory restrictions and was completely replaced by a new VM, called ART (Android Runtime) in Android Lollipop⁶. Security is achieved through application isolation performed by the use of VMs, along with the Linux access control mechanism [11]. Thus, each application is considered as a single user, and has no permission to affect other applications, the operating system, or the user. However, flaws in this model allow privilege escalation attacks [5], bypassing this security model.

The base of the Android stack is the Linux-based *Android Kernel*. This kernel includes hardware abstractions to allow the interaction with the device hardware. It provides the following services: memory management, power management, device drivers, process management, networking and security [12]. Even with the implementation of several security controls, there are vulnerabilities found in this layer. For example, Armando et al. [11] reported a Zygote Process Vulnerability. This vulnerability was fixed by Google, but we discovered that the patch applied works only with Java applications and the vulnerability still reproduce when native code is executed. We do not describe this vulnerability in this paper, because we

²<http://play.google.com/>

³<http://us.lgworld.com/web.main.dev>

⁴<http://www.1mobile.com/>

⁵<http://freakattack.com>

⁶<https://goo.gl/1Fu2s0>

focus on vulnerabilities fixed by us and that vulnerability has not been fixed yet.

The Android operating system evolved throughout the years and, currently, it is in version six, called Marshmallow. In each version, new functionalities are incorporated⁷ and vulnerabilities are corrected [3]. In Android Marshmallow, for example, the permission system was updated, allowing applications to request permissions at run-time, and users to revoke them at any time⁸. Thus, although Android Security has been improved over the years, there are still security breaches available to be explored. Although Android Marshmallow is the newest version, there was no device available with this version during the development of this research. So, this work is based on the previous versions: Lollipop and KitKat.

III. RELATED WORK

In the literature, there are some studies about Android Security. Most of them focus on applications analysis to identify malicious applications [13], [14], [15], [16], [17], [18] or vulnerabilities in legitimate applications [10], [19], [20], [21], [22], [23]. For example, Chin et al. [10] proposed a tool called ComDroid that searches for vulnerabilities in legitimate applications related to Unauthorized Intent Receipt or Intent Spoofing. At the end, the authors presented recommendations for developers to avoid these vulnerabilities.

Zhou and Jiang [17] analyzed four mobile security softwares and identified a malware detection rate between 20.2% and 79.6%. Besides, the authors make available an Android Malware Genome Project⁹. Zhang et al. [16] used these malwares samples and automatically detected 93% of them, using dependency graphs. Likewise, Zhou et al. [18] found 211 malwares in official and alternative Android markets. In addition to the design of an infrastructure for malware detection, Delosières and García [13] emphasize the existence of vulnerabilities in the Android Software Stack and malwares exploiting these vulnerabilities.

Researchers look for this type of vulnerability and propose countermeasures to mitigate them [4], [5], [24], [25], [6], [7], [26]. For example, Chin and Wagner [4] explore two vulnerabilities in WebView and present a tool to detect vulnerable applications. Currently, WebView can be updated from Google Play in the newest Android versions (Android 5.0 and higher) to ensure the bug fixes without OEMs (Original Equipment Manufacturers) dependence¹⁰.

Hei et al. [5] identified two vulnerabilities in Tegra 2 CPU driver and proposed a patch that was accepted by Google. Jang et al. [24] analyzed the accessibility libraries of four computing platforms, including Android itself. They discovered that is possible, for example, to bypass the voice authentication with a replay attack. Moreover, the authors were concerned in identifying the root causes of the vulnerabilities and recommendations for mitigation.

Beyond these related works, there were found vulnerabilities in AAA (Authentication, Authorization and Accounting) protocols of the 3G and 4G networks [6]. It was presented the possibility of using free network services and performing a stealth spam attack, which increases the consumption data [25]. Smalley and Craig [7] showed the use of SELinux to mitigate Android vulnerabilities that resulted in, for example, privilege escalation attacks. It is important to highlight that SELinux is included in the new Android versions¹¹.

Other interesting study was performed by Vidas et al. [27]. Although they did not examine specific vulnerabilities, they analyzed the Android security model focusing on the permission system. Besides, they classified attacks that aim to gain privileged access in accordance with the attacker capabilities. Lastly, they proposed six mitigations to this type of attack. Sun et al. [28] present an analysis of the rooting methods and applications for rooting detection. It is possible to use a rooting detection application to verify if a device is rooted and disable root functions. However, the authors identified the current methods for rooting detections are ineffective. So it is necessary to protect the devices against the rooting methods. In our research, we analyse several types of vulnerabilities, including the exploited by rooting methods, and propose mitigations to fix them.

In the industry side, some IT security companies produce reports that include mobile threats [29], [30], [3], [31], [1], [32]. Often, these reports focus on malware analysis. However, McAfee Labs [1] show that mobile users are still exposed to SSL/TLS vulnerabilities and Google [3] presents vulnerabilities fixed in 2014. Besides, IBM X-Force Team reported some vulnerabilities, such as the Android vulnerability described by Hay [33] that was patched in Android KitKat.

The novelty presented in our paper is the development of mitigation actions of four vulnerabilities, presented in Section V, and the analysis of several vulnerabilities, in accordance with a process that can be replicated, presented in the next section. Additionally one vulnerability discovered by us is presented in subsection V-A, along with its correction actions. Our work follows a similar approach of the penetration test report [34], with the reproduction of vulnerabilities, the evaluation of the impact in the system security, and the proposal of mitigation actions. Thus, the system security can be improved, devising defense strategies. To the best of our knowledge, none of the studies found considered such a vast number of vulnerabilities and analyze all Android software stack. Most focused on malwares or only on one vulnerability.

IV. VULNERABILITIES ANALYSIS

In this paper, we describe our experience in the analysis and mitigation of Android vulnerabilities. Here, results of eight months of work are summarized. A total of 105 vulnerabilities were analyzed, and the results were documented in twelve Attacks Trends Reports (ATRs)¹² prepared between September of 2014 and April 2015.

¹¹<http://source.android.com/devices/tech/security/selinux/>

¹²Please contact the authors for more information on the ATRs.

⁷<http://www.cnet.com/news/history-of-android/>

⁸<http://developer.android.com/guide/topics/security/permissions.html>

⁹<http://www.malgenomeproject.org>

¹⁰<https://goo.gl/x6pB4H>

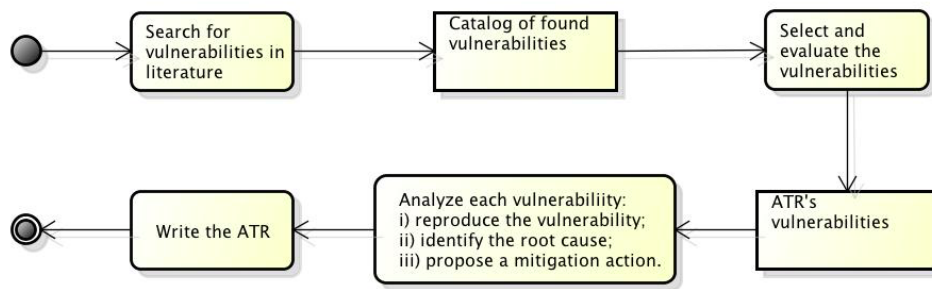


Fig. 2. Analysis Process

In this section, firstly, a process defined by us for the ATR production and the analysis of each vulnerability is described. Next, an overview of the ATRs is done, with the evaluation of the analyzed vulnerabilities.

A. Analysis Process

The analysis process, presented in Figure 2, is proposed to be used to guide the entire team during the ATRs production. The first step is to look for vulnerabilities into specialized sites in the Internet, especially into vulnerabilities databases, and security blogs or forums. The National Vulnerability Database (NVD)¹³, XDA Developers Forum¹⁴ and Kaspersky Blog¹⁵ were some sources used.

The found vulnerabilities are organized in a spreadsheet, called vulnerability catalog, which is filled with basic information extracted from the references. The basic information includes: description, references, reported OS version and device model, and affected layer of the Android stack. Then, a team meeting is carried out to evaluate the vulnerabilities, and decide what vulnerabilities should be analyzed, considering the following aspects:

- the vulnerability explores design features of the Android Software Stack;
- there is enough information to permit the vulnerability reproduction;
- the vulnerability is new and can be reproduced in the newest Android version; and
- the vulnerability is substantially different from the previous analyzed vulnerabilities.

The next step is to try to reproduce these vulnerabilities in flagship devices (top-level device from a specific manufacturer) with the newest and official Android versions. During our vulnerability analysis, the following devices were used: Google Nexus 5, Samsung Galaxy S3, Samsung Galaxy S5, LG G2 and LG G3. It is important to highlight that these devices cover the most important devices of the vendors that dominate the market in the time of this research.

For this step, we search for artifacts used to reproduce each vulnerability, such as Proofs of Concept (PoCs), APKs,

and exploits. In our work, sometimes none of these artifacts were available and our team had to develop a PoC. Another action to try to reproduce the vulnerability is to contact the vulnerability discoverer trying to obtain more information. Vulnerabilities that are reproduced are described using a step-by-step reproduction guide, along with a discussion on the vulnerability exploration.

The next step is a detailed analysis, identifying the root cause of a vulnerability. This analysis is performed considering the particularities of each vulnerability and based on the Android security model. We used, in this step, the two techniques normally used in the literature: dynamic analysis [4], and static analysis [5].

The dynamic analysis is the main mechanism used to verify the vulnerability behavior. Here, the results of the artifact execution were analyzed, collecting logs and using auxiliary tools (e.g. Hooker¹⁶). This analysis is performed in controlled conditions to avoid loss and leakage of sensitive data and, in the end, devices were repaired to the original state.

In the static analysis of a vulnerability, the source code of the artifact is analyzed to identify how a vulnerability is exploited. When the code exploring a vulnerability is unavailable, reverse engineering using existing decompilers (e.g. dex2jar¹⁷) was executed.

After this vulnerability analysis, mitigation actions to correct the vulnerability are proposed. Since these actions depend on the analyzed vulnerability, no generic defense mechanism is defined in the process. In our work, mitigation actions for all analyzed vulnerabilities were proposed, and four vulnerabilities were fixed and described in Section V.

B. Overview of the ATRs

Our industry partner requested the presentation of Android vulnerabilities in an *easy-to-read* report of the analysis results. The ATR template is based on a penetration test report published by Offensive Security [34] and it includes a model for vulnerability evaluation. The partner team validated the ATR template and the evaluation model.

The vulnerability evaluation is based in existing systems, such as Common Vulnerability Scoring System (CVSS)¹⁸

¹³<https://nvd.nist.gov/>

¹⁴<http://forum.xda-developers.com/>

¹⁵<http://blog.kaspersky.com/>

¹⁶<https://github.com/kanpol/hk>

¹⁷<http://sourceforge.net/projects/dex2jar/>

¹⁸<https://nvd.nist.gov/cvss.cfm>

and two metrics are used in the evaluation: *user impact* and *comprehensiveness*. Four levels are defined for each metric: Critical, High, Medium and Low. These metrics are subjective and the evaluation is based in the experience on Android security of our team and in the references of each vulnerability.

The user impact represents the degree with which a vulnerability affects the user. The user impact evaluation is based on each security property: confidentiality, integrity and availability. In the confidentiality analysis, for example, the size of the data leakage and the data sensitiveness (e.g. passwords) are considered. The availability analysis considers how an attack affects the ability of the user to use his/her device. For instance, with which frequency the system is slow or the services/applications are unavailable. In critical cases, the device does not power on (e.g. it enters in boot-loop), where only a factory reset can repair the device. In this case, the malware is deleted, but the user loses his/her data as well. Root vulnerabilities (see subsections V-B and V-C) affect the system integrity because they modify the original system. The creation or modification of unauthorized data are other cases that affect system integrity.

The comprehensiveness measures if a vulnerability is widely exploited, affecting several Android versions, vendors, countries and user groups. For example, the VPN encryption keys vulnerability, described in subsection V-A, affects only the rooted devices. Additionally, the comprehensiveness evaluation considers if a vulnerability is easily exploited and disseminated. This analysis verifies also if the user can identify an eventual attack, and if an user can easily protect or repair his/her device.

Each ATR contains the following sections: *Executive Summary*; *Summary of Results*; *Vulnerability Analysis Template*; and *Vulnerabilities Analysis*. The *Executive Summary* shows basic definitions, including those related to the vulnerability evaluation.

The Summary of Results of an ATR presents a brief description of the selected vulnerabilities for this ATR, with an evaluation of the user impact and comprehensiveness. Figure 3 exposes the evaluation of the user impact and comprehensiveness of the analyzed vulnerabilities. It is possible to verify the importance of treating the Android security flaws, since 42,86% of the analyzed vulnerabilities has high or critical impact and high or critical comprehensiveness.

Figure 4 presents the amount of vulnerabilities in accordance with the security property and the software layer affected by the vulnerability. It is important to highlight that one vulnerability can affect more than one security property. Despite the attempts to look for vulnerabilities in all Android Software Layers, the found vulnerabilities affect mainly the Application and Framework layers. A reasonable number of integrity vulnerabilities in the Kernel layer can be observed, related mostly to root vulnerabilities.

One of the goals of our process is to verify if the vulnerabilities are still being reproduced. In that direction, we were able to reproduce 45 vulnerabilities (42.86%) in at least one device. 35 vulnerabilities (36.46%) were reproduced in LG

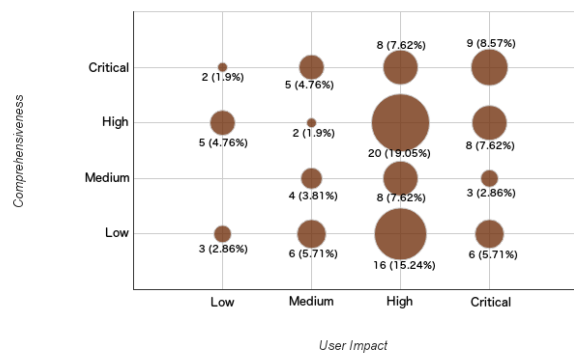


Fig. 3. User Impact and Comprehensiveness

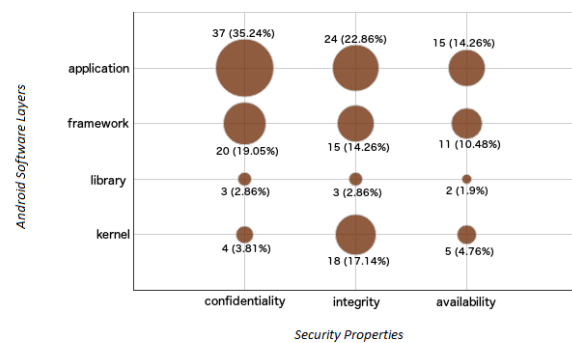


Fig. 4. Security Properties and Android Software Layers

devices (G2 and G3), where 20 vulnerabilities (20.83%) in Nexus 5 and 28 vulnerabilities (29.17%) in other devices (e.g. Galaxy S5 and S3). However, 19 vulnerabilities (18.1%) were fixed and do not reproduce anymore in our devices.

Unfortunately, not all selected vulnerabilities could be tested, mainly because of the following reasons: i) lack of artifacts necessary to perform the reproduction (e.g. specific phone model); ii) lack of enough information to verify if a vulnerability is being explored (e.g. detection of a data leakage). So, 41 vulnerabilities (39.05%) were not tested in our devices.

The report of each vulnerability analysis is written in accordance with the Vulnerability Analysis Template. Besides the basic information, a ATR contains the step-by-step guide for reproduction and the analysis results with the identification of the vulnerability root cause and the suggested mitigation actions. We also observed the particularities in each vulnerability that hinder the creation of a generic solution to fix all vulnerabilities. In the next section, we detail the analysis and mitigation of a subset of four vulnerabilities.

V. MITIGATED VULNERABILITIES

The initial focus of this research was to study Android existing/community-reported vulnerabilities, however, we were also able to propose the process described in the previous section and identify an unreported vulnerability.

This section reports our experience on analysing, reproducing and correcting four vulnerabilities: i) VPN Encryption Keys Vulnerability, ii) IORoot, iii) New Root Method, and iv) Cross-Signed Certificate. For each vulnerability, the analysis process is detailed, including a description of the correction applied.

A. VPN Encryption Keys Vulnerability

The VPN Encryption Keys vulnerability is the non-cataloged vulnerability detected by us. It is important to stress that this vulnerability is restricted to IPSec (IP Security Protocol). This vulnerability allows malicious apps to access the cryptographic keys used in a Virtual Private Network (VPN) connection and permits this information to be sent to third parties. A malicious app is also able to log all data transmitted over the active VPN connection. With these two information, it is possible to compromise the secrecy of communication, deciphering VPNs packages with applications such as Wireshark¹⁹.

The user impact of this vulnerability was considered *High* due to the possibility of leaking information transmitted through a VPN. On the other hand, the comprehensiveness was considered *Medium*, since it only works on rooted devices, even though it affects all android models and versions. However, it is possible an dual attack, rooting a device before exploiting the VPN vulnerability. This is a network vulnerability that affects the *Kernel* layer from the Android Architecture.

Our team performed the analysis and reproduction of this vulnerability as follows. A PoC was developed with two buttons: one to start; and one to finish the packet capturing process. Before beginning capturing the packets, the user must manually perform the connection to a VPN. The packets are captured using an open source packet sniffer, TCPDump²⁰, which was embedded in the PoC. While the VPN and sniffer are activated, we access some HTTP site to be verified in deciphered packets, facilitating the analysis process. After finishing the capture of the packets, the TCPDump saves the data in a .pcap file. While still connected to the VPN, the PoC also executes the command `ip xfrm state` that returns the VPN connection cryptographic keys that are saved in other file.

Afterwards, with the help of a PC, it is possible to open the .pcap file on Wireshark and enable ESP (Encapsulating Security Payload) decipher. The other file contains the information of the two Security Associations (SAs) created while establishing the IPSec based VPN connection. In Wireshark, the SA ESPs are configured, as indicated in Figure 5. If all information was provided correctly, Wireshark decipheres the data and display the deciphered packets.

The PoC was tested in LG G2, LG G3 and Nexus 5. The vulnerability was successfully reproduced in all devices and versions (Android 4.2.2 and 5). It is important to highlight the fact that any malicious app could perform these steps,

automatically, through a background service. It is possible to verify if the device is connected to a VPN, starting the capture of packets when the VPN connection is identified. The cryptographic keys are also captured and these information are sent by e-mail or to a third party server.

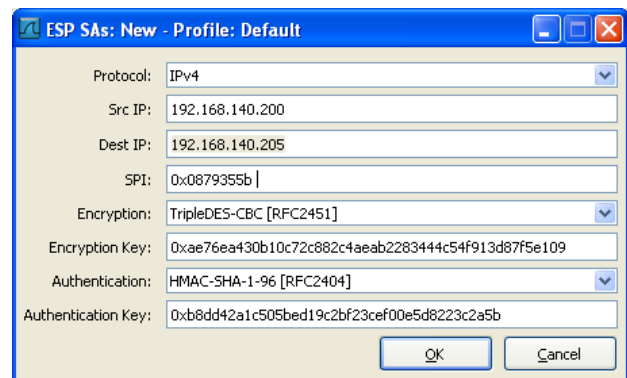


Fig. 5. ESP SA Example

In principle, it is not possible to avoid the capture of packets, so we should protect the VPN encryption keys. So, we considered the `ip` command as the root cause of the vulnerability. Then, we suggested as a mitigation action the removal of the daemon `ip` or `xfrm` parameter of this command. To obtain a more effective solution, it is necessary to protect the devices against root methods, because it is possible to restore this command in a rooted device.

The fix was implemented in LG G3 with the removal of the `ip` daemon. This fix was applied only in LG G3, because we only have the G3 code. Further tests were performed to verify that this vulnerability does not reproduce anymore and that none Android service is affected. We considered this solution satisfactory since the device worked as intended, even when connected to a VPN, and because we did not find in the literature references that prove the necessity of this command. This command, for example, does not exist in the MAC OS. However, it is important to highlight that this command, and consequently the vulnerability, still exists in other Unix-based systems.

B. IORoot

A rooted device allows to exploit vulnerabilities, such as the VPN vulnerability described in previous subsection. Besides, rooting the device is not recommended by the OEMs, pointing out that the warranty of the device is canceled after the user performs this procedure. Although OEMs try to protect their devices against rooting procedures, there are wide available mechanisms/apps to perform device rooting. While some of these mechanisms target specific kernel/OS versions, like KingRoot²¹ and TowelRoot²², others are developed for a specific device/OEM, such as IORoot and New Root Method, targeting LG flagship devices. The analysis of the IORoot²³ is

¹⁹<https://www.wireshark.org/>

²⁰<http://www.tcpdump.org/>

²¹<http://www.kingroot.net>

²²<http://towelroot.info>

²³<http://goo.gl/jyY8Ft>

described here, while the New Root Method is detailed in the next subsection.

Root vulnerabilities affect the Android kernel and make the device vulnerable for allowing applications to access and modify device sensitive information, then breaking its integrity. The rooting methods are a big concern for OEMs, and they are always looking for protective mechanisms against these methods [28]. For these reasons, the user impact was considered *High*. Despite IORoot affects many Android versions, it only affects devices from one OEM and the user him/herself must perform the rooting process. Thus, the comprehensiveness was evaluated as *Medium*.

IORoot was originally developed as a Windows script to root the LG G2 on 4.2.2. Since then, IORoot has been updated and its script now supports several other LG devices. It was based on the ADB sideload²⁴, which permits the user to execute scripts as a device administrator. To obtain root privileges on a device through this method, it is necessary to connect the device to a PC using a USB cable, turn on the device debug mode and execute the file `ioroot.bat` on Windows or `ioroot.sh` on Linux. Thereby, an update patch, named `datroot.zip`, is installed. It is important to highlight that the PC should have pre-installed adb-tools. This vulnerability was tested and successfully reproduced in LG G2 and G3. However, this vulnerability does not reproduce in G3 with Android Lollipop.

In the analysis process of this vulnerability, we observed that the patch basically contains `su` and `daemonsu` binaries, and the `install-recovery.sh` script. These files are extracted and copied to the target device. An important fact is that the root privilege is obtained when the `install-recovery.sh` script is executed as device admin and the `daemonsu` process is initiated. The `su` binary allows any app to execute as root. Unfortunately, during sideload update process, a patch signature verification was supposed to be performed, but a failure in this step permitted any script to be executed as device admin. This verification is fundamental to guarantee that only OEMs' scripts should be accepted, so, this verification is considered the root cause of this vulnerability. We performed a dynamic analysis to understand the update process by sideload, and we found the functions executed in this process. We analyzed these functions and checked that the signature verification is not performed.

In this context, as a mitigation action, we suggested to modify the sideload, accepting only files signed by the OEM, or to remove the sideload feature. We decided to remove the sideload, because is not an indispensable feature for OEM devices. This feature is mostly used to update Android OS and there are other methods more adequate (e.g. Firmware Over the Air). To remove the sideload, we analyzed the functions used when the update by adb sideload is chosen from recovery menu. We found the `service_to_fd` function of `/bootable/recovery/minadb/services.c` that starts the sideload service. We commented the lines inherent to this service initialization, as shown in Figure 6, and an error is returned

when this option is requested. This approach was successfully applied in LG G3, in such a way that this rooting method stopped working. When the G3 Lollipop source code was analyzed, the IORoot vulnerability does not reproduce, because the patch signature is verified.

```
int service_to_fd(const char *name) {
    int ret = -1;
    /*
        if (!strcmp(name, "sideload:", 9)) {
            ret = create_service_thread(sideload_service,
                (void*) atoi(name + 9));
            #if 0
        } else if (!strcmp(name, "echo:", 5)) {
            ret = create_service_thread(echo_service, 0);
            #endif
        }
    */
    if (ret >= 0) {
        close_on_exec(ret);
    }
    return ret;
}
```

Fig. 6. Snippets codes for disabling the sideload service

C. New Root Method

The New Root Method is another root vulnerability analyzed in this paper and it allows the user to gain root access after the execution of an update script. This method is comprised of a series of steps detailed by XDA Developers [35]. It uses the *DownloadMode*, which is a mechanism for flashing radio firmware/ROM upgrade. To be able to communicate with the device under DownloadMode, it uses an Windows executable file called `Send_Command.exe`. It connects with the device using a USB interface and sends commands to DownloadMode. The main command sent to device is to execute the script `lg_root.sh`. Basically, among other things, this script puts the files normally used to root a device (`su`, `daemonsu` and `install-recovery.sh`), sets appropriate permissions and installs *SuperSu* app.

The user impact and the comprehensiveness was evaluated as *High* and *Medium*, respectively, similarly to the IORoot vulnerability. In our tests, LG G2 and LG G3 were successfully rooted using Android KitKat. An update on this vulnerability is already available for Lollipop devices²⁵.

Since the `Send_Command's` code was not available, the mitigation action proposed is based on the knowledge acquired by studying the code of DownloadMode. The strategy was to monitor the steps to perform this root method by placing log messages during the execution of each step, with a posterior analysis of these steps. We observed that is necessary to avoid the execution of the `lg_root.sh` script by `sh` process. Then, one possible action was to modify the permissions of `sh` process, which was found to not solve the issue.

Another solution was to verify if the execution of `sh` command is requested by DownloadMode and to block it. We thought about blocking the `lg_root.sh` script, but it is possible

²⁴<http://goo.gl/yU4o1z>

²⁵<http://goo.gl/d7LFh7>

to rename the file name to bypass this approach. Thus, to prevent the New Root Method vulnerability to be executed, a verification was added into the *laf_cmd_execute_proc* function of *laf_cmd.c* source code. This function executes the commands sent to the device, and we added a test to run only commands different from *sh*, as shown in Figure 7. This approach was successfully applied, in such a way that this rooting method stopped working.

```

...
switch ((pid = fork())) {
case -1:
    LAF_LOGE("Fork failed, errno=%d\n", errno);
    ret = LAF_ERROR_REQUEST_ABORTED;
    break;
case 0: /* child */
    LAF_LOGD("child pid = %ld\n", (long)getpid());
    close(pipefd[0]); /* close unused read end */
    dup2(pipefd[1], STDOUT_FILENO);
    close(pipefd[1]);
    // Block sh process
    if(strcmp(argv[0], "sh") == 0) {
        ret = -1;
    } else {
        ret = execvp(argv[0], argv);
    }
    if(ret == -1) LAF_LOGE("execvp failed. error = %d\n", errno);
    exit(0);
default: /* parent */
    close(pipefd[1]); /* close unused write end */
    while (read(pipefd[0], ss->data[1en], 1) > 0) {
        s->data;
    }
}
...

```

Fig. 7. Snippets codes of the *sh* process blockade

D. Cross-Signed Certificate

The Cross-Signed Certificates Vulnerability was reported by the security company Trend Micro²⁶. These certificates can be created when an entity A signs the certificate of an entity B, while at the same time, an entity B signs the certificate of an entity A, as shown in Figure 8. Thus, two cross-signed certificates result in a looped certificate chain. Unfortunately, the Android OS does not correctly handle these certificates and the system can have unexpected behavior, such as system slow down or forced reboot. There were two identified attacks scenarios: import a malformed PKCS#12 file with a loop certificate chain into Android, and install an app signed by one of the cross-signed certificates.



Fig. 8. Cross-signed certificates examples

This vulnerability affects only the system availability with a temporary system slow down and the system is back to work normally after sometime. In our tests, we detected that some

²⁶<http://goo.gl/QtORhY>

functions (e.g. install an app) only work after a system reboot. Thus, we considered that the user impact is *Low*. It is possible to develop a malware that executes an attack during the device booting, rising the user impact since it becomes necessary a factory reset. Although we observed that the vulnerability affects all pre-5.1 Android versions, the comprehensiveness was considered *Medium* because we did not find any return to the attacker and the system is easily repaired.

The first step to reproduce this vulnerability is the creation of two cross-signed certificates, using *keytool* and *openssl* commands. Henceforward, to simplify the description of the process to create the certificates, the entities A and B will be called Alice and Bob respectively. Firstly, a JKS keystore is generated with an Alice self-signed certificate. This certificate is exported to p12 format and then Alice's private key is extracted. With *openssl* commands, Bob's self-signed certificate is created. After creating the CSRs (Certificate Signing Request) to Alice and Bob, an Alice certificate signed by Bob and a Bob certificate signed by Alice are generated. The next step it the manual edition of the Alice certificate file, adding the Bob certificate. Finally, the keystore is updated with Bob's self-signed certificate and the new Alice certificate, replacing Alice's self-signed certificate.

In the first attack scenario, the cross-signed certificates are exported to p12 file that is imported on the device. In the other scenario, an app is signed with the created keystore before being installed. We reproduced both scenarios in all tested devices and Android versions, except on Nexus with Android 5.1. The failures in the certificate validation by Android Framework allow the attacks because there is no loop identification in the certificate chain. For example, in the installing app scenario, the *createChain* method from *JarUtils* class only verify if there is a next certificate in the chain and if it is self-signed.

We fixed this validation process, adding a test in the *createChain* method to verify if the next certificate in the chain has not been previously verified. If a cross-signed certificate is found, a *SecurityException* is launched. Then, it was necessary also to add an exception treatment in *onCreate* method from *PackageInstallerActivity* class. These corrections can be seen in Figure 9.

For another scenario (import a malformed PKCS#12 file), the solution is a little different. We also create a list with the whole certificate chain, but if a cross-signed certificate is found, we return *null* to avoid the installation of malformed certificates. This modification was done into *engineGetCertificateChain* method from *PKCS12KeyStoreSpi* class. Due to this modification, a *NullPointerException* can be triggered in *extractPkcs12* method from *CredentialHelper* class. However this exception should be treated posteriorly, in *doInBackground* method from *CertInstaller* class. This method is an *AsyncTask* responsible for the background installation of the certificates, and we modified it to indicate that the PKCS#12 file is invalid when cross-signed certificates found. In the failure case, the system thinks that the password is invalid and ask it again. So, lastly, it was necessary to add a test


```

private static X509Certificate[] createChain(X509Certificate signer,
X509Certificate[] candidates, boolean chainCheck) {
    LinkedList chain = new LinkedList();
    ...
    while (true) {
        issuerCert = findCert(issuer, candidates, subjectCert, chainCheck);
        if (issuerCert == null) {break;}
        chain.add(issuerCert);
        ...
        if (issuers.contains(issuerCert.getSubjectDN())) {
            throw new SecurityException("Cross-signed Certificate");
        }
        ...
    }
    return (X509Certificate[])chain.toArray(new X509Certificate[count]);
}

protected void onCreate(Bundle icicle) {
    ...
    try {
        parsed = PackageUtil.getPackageInfo(sourceFile);
    } catch (Exception exception) {
        Log.w(TAG, "Failed to obtain package info.");
        showDialogInner(DLG_PACKAGE_ERROR);
        setPmResult(PackageManager.INSTALL_FAILED_INVALID_APK);
        mInstallFlowAnalytics.setPackageInfoObtained();
        mInstallFlowAnalytics
            .setFlowFinished(InstallFlowAnalytics.RESULT_FAILED_TO_GET_PACKAGE_INFO);
        return;
    }
    ...
}

```

Fig. 9. Snippets codes of the correction in the app installation scenario

in *onExtractionDone* method from *CertInstaller* class. In this test, if the PKCS#12 file is invalid, the user receives an error message and the import certificate procedure is finished.

It is possible to observe that, in our correction, applications signed by cross-signed certificates are not installed and certificates from a malformed PKCS#12 files are not imported. Additionally, we also created a plan in CTS (Compatibility Test Suit)²⁷ to verify if our solutions works.

VI. CONCLUSIONS AND FUTURE WORK

As the number of malwares and security attacks aimed at mobile devices rises, so does the need to devise defense strategies, in order to protect the user and his/her personal and business information. In that scenario, protection mechanisms tailored to the Android platform become essential, since Android calls for more than 80% of the market share.

In this paper, we reported our experience in the analysis and mitigation of Android vulnerabilities. The high number of existing vulnerabilities and the impact of them in the Android security show the importance of this topic and the necessity of devising defense mechanisms.

Although the analyzed vulnerabilities are known, some of them can still be exploited. Then, we analyzed existing vulnerabilities, performing attacks in Android devices and discussing the attack impact. The goal was to identify attacks that could be reproduced more easily and the attacks with higher impact for the user. Each vulnerability was analyzed individually, using static and dynamic analysis techniques. Thus, it is possible to understand how the vulnerability is exploited in order to perform the attack and then to propose a protective mechanism for each vulnerability. Besides, we improved the Android security with the correction of four vulnerabilities.

In this research, an analysis process was defined and used in the analysis of more than 100 vulnerabilities, summarized in this paper. Four mitigated vulnerabilities were chosen to be

²⁷<https://source.android.com/compatibility/cts/index.html>

detailed in this paper, following this process. In analysis of the root methods and cross-signed certificate vulnerability, we used the dynamic analysis to identify the path taken during an attack and the executed functions. After, we analyzed the source code and proposed code fixes to avoid the attacks.

Moreover, when running that process, another important contribution was the discovery of one unreported vulnerability. This vulnerability allows malicious apps to access the cryptographic keys used in a Virtual Private Network (VPN) connection, to log all data transmitted over the active VPN connection, and permits this information to be sent to third parties to be unencrypted using the collected cryptographic keys. That vulnerability was identified, analyzed and reported, highlighting its root cause and deploying a mitigation action.

As future work, new vulnerabilities can be analyzed and fixed using the proposed process. With the acquired knowledge, new vulnerabilities can be discovered. An interesting research is also the identification of common features of the vulnerabilities to classify them. Thus, mechanisms to analyze and mitigate a specific type of vulnerability, or improvement for Android operating system can be proposed. It is also possible to study the tools used in static and dynamic analysis, proposing improvements or new tools.

Besides, vulnerabilities resulted from application developer negligence can be considered, generating reports similar to those presented here. Then, guidelines and tools for helping developers and the whole development teams to improve the security of their software, along with a development process could be presented.

ACKNOWLEDGMENT

This work was funded by MCT/Informatics Law under project grant number 3076. Carlos André Batista de Carvalho was also supported by CAPES/FAPEPI Doctoral Scholarship, and Rossana Maria de Castro Andrade by CNPq Research Scholarship.

Special thanks to our Security Project Team: Andressa Bezerra Ferreira, Arthur Paulino, Bruno Góes, Carol Magalhães, Creonilson Rodrigues, Édipo Souza, Ítalo Romeiro, Letícia Fernandes, Manoel Fiuza, Paulo Artur Duarte and Roney Gomes. Their contribution was of inestimable importance for the execution and conclusion of this work.

REFERENCES

- [1] McAfee Labs, "McAfee Labs Threats Report: February 2015," <http://goo.gl/tKTzXo>, 2015, accessed: 2015-09-04.
- [2] S. Liebergeld and M. Lange, "Android security, pitfalls and lessons learned," in *Information Sciences and Systems*, ser. Lecture Notes in Electrical Engineering, 2013, vol. 264, pp. 409–417.
- [3] Google, "Google report android security 2014 year in review," <https://goo.gl/N5Sv3>, 2014, accessed: 2015-09-04.
- [4] E. Chin and D. Wagner, "Bifocals: Analyzing webview vulnerabilities in android applications," in *Information Security Applications*, ser. Lecture Notes in Computer Science, 2014, vol. 8267, pp. 138–159.
- [5] X. Hei, X. Du, and S. Lin, "Two vulnerabilities in android os kernel," in *International Conference on Communications (ICC)*, 2013, pp. 6123–6127.
- [6] C. Peng, C.-Y. Li, H. Wang, G.-H. Tu, and S. Lu, "Real threats to your data bills: Security loopholes and defenses in mobile data charging," in *ACM Conference on Computer and Communications Security (CCS)*, 2014, pp. 727–738.

- [7] S. Smalley and R. Craig, "Security Enhanced (SE) android: Bringing flexible mac to android," in *Network and Distributed System Security Symposium (NDSS)*, 2013.
- [8] A. Shabtai, Y. Fledel, U. Kanonov, Y. Elovici, S. Dolev, and C. Glezer, "Google android: A comprehensive security assessment," *IEEE Security and Privacy*, vol. 8, no. 2, pp. 35–44, 2010.
- [9] M.-C. Chen, J.-L. Chen, and T.-W. Chang, "Android/osgi-based vehicular network management system," *Computer Communications*, vol. 34, no. 2, pp. 169–183, 2011, special Issue: Open network service technologies and applications.
- [10] E. Chin, A. P. Felt, K. Greenwood, and D. Wagner, "Analyzing inter-application communication in android," in *International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2011, pp. 239–252.
- [11] A. Armando, A. Merlo, M. Migliardi, and L. Verderame, "Would you mind forking this process? a denial of service attack on android (and some countermeasures)," in *IFIP International Information Security and Privacy Conference (SEC)*, 2012, pp. 13–24.
- [12] S. Gunasekera, *Android Apps Security*, 1st ed. Apress, 2012.
- [13] L. Delosières and D. García, "Infrastructure for detecting android malware," in *Information Sciences and Systems*, ser. Lecture Notes in Electrical Engineering, 2013, vol. 264, pp. 389–399.
- [14] K. Allix, Q. Jerome, T. F. Bissyande, J. Klein, R. State, and Y. L. Traon, "A forensic analysis of android malware – how is malware written and how it could be detected?" in *IEEE 38th Annual Computer Software and Applications Conference (COMPSAC)*, 2014, pp. 384–393.
- [15] Z. Yang, M. Yang, Y. Zhang, G. Gu, P. Ning, and X. S. Wang, "Appintert: Analyzing sensitive data transmission in android for privacy leakage detection," in *ACM Conference on Computer and Communications Security (CCS)*, 2013, pp. 1043–1054.
- [16] M. Zhang, Y. Duan, H. Yin, and Z. Zhao, "Semantics-aware android malware classification using weighted contextual api dependency graphs," in *ACM Conference on Computer and Communications Security (CCS)*, 2014, pp. 1105–1116.
- [17] Y. Zhou and X. Jiang, "Dissecting android malware: Characterization and evolution," in *IEEE Symposium on Security and Privacy (S&P)*, 2012, pp. 95–109.
- [18] Y. Zhou, Z. Wang, W. Zhou, and X. Jiang, "Hey, you, get off of my market: Detecting malicious apps in official and alternative android markets," in *Network and Distributed System Security Symposium (NDSS)*, 2012, pp. 409–417.
- [19] W. Enck, P. Gilbert, B.-G. Chun, L. P. Cox, J. Jung, P. McDaniel, and A. N. Sheth, "Taintdroid: An information-flow tracking system for realtime privacy monitoring on smartphones," in *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2010, pp. 1–6.
- [20] L. Lu, Z. Li, Z. Wu, W. Lee, and G. Jiang, "Chex: Statically vetting android apps for component hijacking vulnerabilities," in *ACM Conference on Computer and Communications Security (CCS)*, 2012, pp. 229–240.
- [21] M. Spreitzenbarth, F. Freiling, F. Echter, T. Schreck, and J. Hoffmann, "Mobile-sandbox: Having a deeper look into android applications," in *Annual ACM Symposium on Applied Computing (SAC)*, 2013, pp. 1808–1815.
- [22] F. Wei, S. Roy, X. Ou, and Robby, "Amandroid: A precise and general inter-component data flow analysis framework for security vetting of android apps," in *ACM Conference on Computer and Communications Security (CCS)*, 2014, pp. 1329–1341.
- [23] Y. Zhang, M. Yang, B. Xu, Z. Yang, G. Gu, P. Ning, X. S. Wang, and B. Zang, "Vetting undesirable behaviors in android apps with permission use analysis," in *ACM Conference on Computer and Communications Security (CCS)*, 2013, pp. 611–622.
- [24] Y. Jang, C. Song, S. P. Chung, T. Wang, and W. Lee, "Allly attacks: Exploiting accessibility in operating systems," in *ACM Conference on Computer and Communications Security (CCS)*, 2014, pp. 103–115.
- [25] C. Peng, C.-Y. Li, G.-H. Tu, S. Lu, and L. Zhang, "Mobile data charging: New attacks and countermeasures," in *ACM Conference on Computer and Communications Security (CCS)*, 2012, pp. 195–204.
- [26] L. Wu, M. Grace, Y. Zhou, C. Wu, and X. Jiang, "The impact of vendor customizations on android security," in *ACM Conference on Computer and Communications Security (CCS)*, 2013, pp. 623–634.
- [27] T. Vidas, D. Votipka, and N. Christin, "All your droid are belong to us: A survey of current android attacks," in *USENIX Workshop on Offensive Technologies (WOOT)*, 2011, pp. 81–90.
- [28] S.-T. Sun, A. Cuadros, and K. Beznosov, "Android rooting: Methods, detection, and evasion," in *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*, ser. SPSM '15, 2015, pp. 3–14.
- [29] Cenizic, "Application vulnerability trends report," 2014.
- [30] F-Secure, "Threat report h2 2014," <https://goo.gl/OMSFO9>, 2014, accessed: 2015-09-04.
- [31] Lookout, "Mobile threat report 2014," <https://goo.gl/bnt10q>, 2014, accessed: 2015-09-04.
- [32] Symantec, "Internet security threat report 2015," <https://goo.gl/aP5L8F>, 2015, accessed: 2015-09-04.
- [33] R. Hay, "Android collapses into fragments," <https://goo.gl/utOXFd>, 2013, accessed: 2015-09-04.
- [34] Offensive Security, "Penetration test report," <https://goo.gl/rBjXLK>, 2013, accessed: 2015-09-04.
- [35] XDA Developers, "New root method for lg devices," <http://goo.gl/MVGhj6>, 2015, accessed: 2015-09-04.

Performance Analysis of Proposed Network Architecture: OpenFlow vs. Traditional Network

Idris Z. Bholebawa^{#1}, Rakesh Kumar Jha^{*2}, Upena D. Dalal^{#3}

[#]Department of Electronics and Communication Engineering, S. V. National Institute of Technology, Surat, Gujarat, India. 395007

^{*}School of Electronics and Communication Engineering, Shri Mata Vaishno Devi University, Katra, J&K 182320

Abstract – The Internet has been grown up rapidly and supports variety of applications on basis of user demands. Due to emerging technological trends in networking, more users are becoming part of a digital society, this will ultimately increases their demands in diverse ways. Moreover, traditional IP-based networks are complex and somehow difficult to manage because of vertical integration problem of network core devices. Many research projects are under deployment in this particular area by network engineers to overcome difficulties of traditional network architecture and to fulfill user requirements efficiently. A recent and most popular network architecture proposed is Software-Defined Networks (SDN). A purpose of SDN is to control data flows centrally by decoupling control plane and data plane from network core devices. This will eliminate the difficulty of vertical integration in traditional networks and makes the network programmable. A most successful deployment of SDN is OpenFlow-enabled networks.

In this paper, a comparative performance analysis between traditional network and OpenFlow-enabled network is done. A performance analysis for basic and proposed network topologies is done by comparing round-trip propagation delay between end nodes and maximum obtained throughput between nodes in traditional and OpenFlow-enabled network environment. A small campus network have been proposed and performance comparison between traditional network and OpenFlow-enabled network is done in later part of this paper. An OpenFlow-enabled campus network is proposed by interfacing virtual node of virtually created OpenFlow network with real nodes available in campus network. An implementation of all the OpenFlow-enabled network topologies and a proposed OpenFlow-enabled campus network is done using open source network simulator and emulator called Mininet. All the traditional network topologies are designed and analyzed using NS2 - network simulator.

Keywords – SDN, OpenFlow, Mininet, Network Topologies, Interfacing Network.

I. INTRODUCTION

Today the Internet has vastly developed and modern technologies can be easily deployed based on networking standards. This leads to a huge community of digital society accessible at anytime and from anywhere. The Internet has been grown up rapidly and supports variety of applications that

includes web services, broadcasting multimedia, database accessing, entertainments, real-time audiovisual communications (AVC), computer-to-computer communication, etc. This traditional network architecture is distributed control and the elements used for the particular complex design is controlled by its own firmware installed in their storage space. However, in spite of this well-developed approach, a traditional IP-based networks are still complex and are somehow difficult to manage. A network complexity metrics and network management related issues is discussed in [1]. In order to deploy any desire network policies in a bounded network architecture, because of its distributed nature, a network administrator requires to configure each and individual network devices separately. This will be time consuming and may lead to increase complexity or may introduce faults due to reconfiguration of individual hardware or due to load unbalance. The challenging task for network architecture is that the network requires to be dynamic in nature and can be controlled automatically without involvement of manual operator. This type of automatic reconfiguration of hardware and management task is deficient in traditional IP-based network. Moreover, in traditional network the core networking devices are vertically integrated. The term vertically integrated is mentioned in [2], because the controlling element to handle network traffic and the forwarding element to forward network traffic, in form of digital packets as directed by control element, is bundled inside a common networking device. This in turn reduces flexibility in designing related matter and made it to endure as a static architecture. Thus, it can ultimately confine innovative ideas related to evolution and deployment of bounded infrastructure network.

To overcome the difficulties encountered in traditional IP-based network architecture and its underlying technologies, a new area of research challenges the evolution of Next Generation Networks (NGN) [3]. Numerous approaches are taken into consideration for the efficient deployment of Future Internet [4]. Various research topics and research projects has already been going on under government as well as under private entities as discussed in [5]. Out of the several contribution associated in order to improvement of rigid traditional infrastructure network, one of the recent approach for emerging networking paradigm is

the Software Defined Networking (SDN) [6, 7]. SDN is a kind of programmable network that gives hope to overcome the limitations faced by current traditional network infrastructure in diverse ways. Firstly, it decouples the control plane from the underlying network core devices like routers and switches that forwards data traffic (also called data plane), this in turn will breaks the vertical integration difficulty problem. Secondly, with the separation of data plane and control plane, all the controlling task is handled by a centralized software-based controller and the core devices are only responsible for data traffic forwarding task as directed by logically centralized controller. In SDN, the controller is a sort of controlling element that controls underlying networking devices through well-developed programmable coding scheme. Due to this configuration, the controller as well as the software running inside the controller is also referred as Network Operating System (NOS) as discussed in [8]. A traditional network architecture and a simplified SDN network architecture differentiating control and data plane is shown in fig. 1 and fig. 2 respectively.

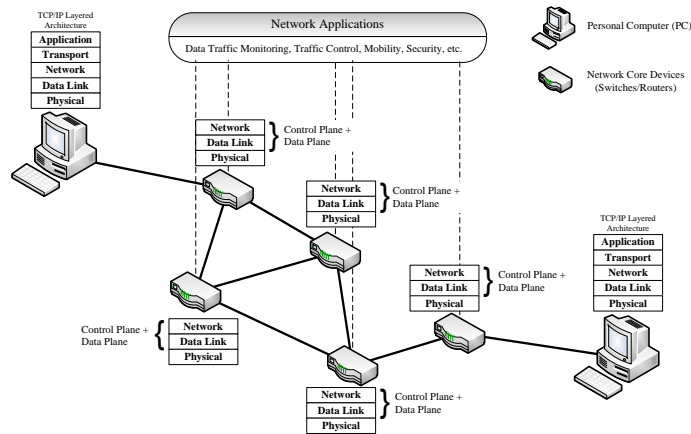


Fig. 1 Simplified Traditional Network Architecture

To realize the above discussed SDN architecture a protocol was standardized by Open Networking Foundation (ONF) in 2011 [9]. The first standard of SDN defined by ONF is OpenFlow networking protocols [10]. OpenFlow was initially deployed at Stanford University as a clean slate project by N. McKeown et.al in 2008 [11]. Today, more than 20 firms has deployed OpenFlow networks, including universities and giant networking companies [9]. OpenFlow networks have specific capabilities of controlling multiple hardware switches (also called OpenFlow-enabled switches) by means of a single centrally controlled logic (also called OpenFlow controller) on a secure channel with the help of OpenFlow Protocols (OFP) [12] as shown in fig. 2.

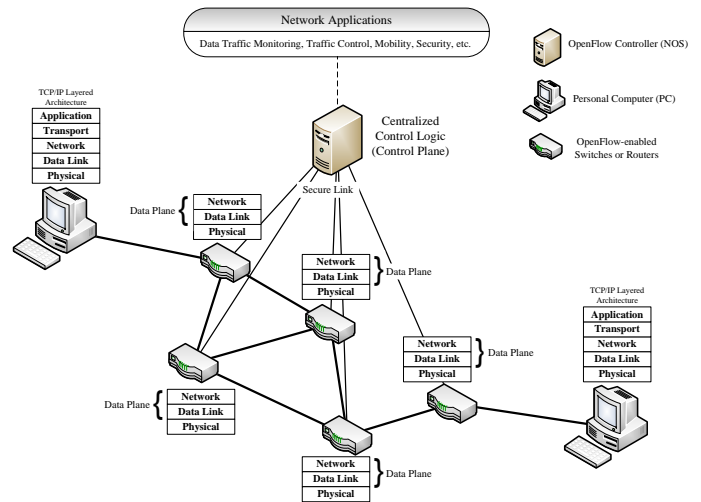


Fig. 2 Simplified Software-defined Network Architecture

In this paper a comparative performance analysis of above discussed traditional IP-based network with OpenFlow-based network is done. A performance analysis is done by implementing different network topologies and comparing the results obtained on execution of proposed networks. An OpenFlow-enabled campus network architecture is also proposed in this paper and performance comparison of this proposed network with traditional network is analyzed. A proposed campus network architecture is designed by interfacing virtual SDN nodes with real infrastructure nodes. The network architecture is proposed using an optimized open-source simulator (NS-2 and Mininet). The paper is organized by discussing a brief introductory concepts of SDN and OpenFlow networks and its architectural design and functioning in Section II. A topological comparison between traditional and OpenFlow-based network and result analysis for the same is done in Section III, the traditional and OpenFlow-based network topologies are designed using an open source simulation tool discussed in the same section. An effective campus network architecture has been proposed using the same open source tool, in that a virtual nodes are interfacing with real nodes is enlightened in Section IV and a comparative performance analysis between traditional and OpenFlow-enabled network for the same proposed architecture is done. Lastly, we conclude the topic in Section V on the basis of results obtained and a brief idea related to future perspective is deliberated.

II. SOFTWARE-DEFINED NETWORKS: OPENFLOW-BASED NETWORK ARCHITECTURE

SDN differs from traditional network architecture with the separation of data plane and control plane as discussed above. SDN comes into existence due to the deployment of OpenFlow technology at Stanford University [11]. Initially OpenFlow networks was deployed as a campus network, then in 2011 ONF

started a working group of SDN through open standard deployments by grouping several giant networking companies [9]. Currently ONF is promoting both SDN and OpenFlow, and standardize OpenFlow as a first SDN standard for software-defined architecture. OpenFlow-based network architecture comprises of three layers as shown in fig. 2. The lowest layer is data plane, which is only responsible for forwarding data packets coming towards its ingress port, the second layer is control plane, which resides above data plane and control all the underlying data forwarding element centrally by means of centralized control logic. The third and uppermost layer is application plane, this plane is hypothetically available. Generally all the forwarding element is controlled by a single controller, the role of application plane is to define practical applications required for efficient data traffic flow at the beginning of network execution. This layer is solely responsible for management and data forwarding related tasks.

The communication between data plane and control plane is done on a secure link by means of OFP as discussed above. OFP supports three message types, *controller-to-switch*, *asynchronous and symmetric*, a detailed description of OFP is given in [12]. The OpenFlow-enabled switch comprising data plane consists of flow tables (or look-up tables) for data flow entries. A components required for an OpenFlow network architecture is shown in fig. 3.

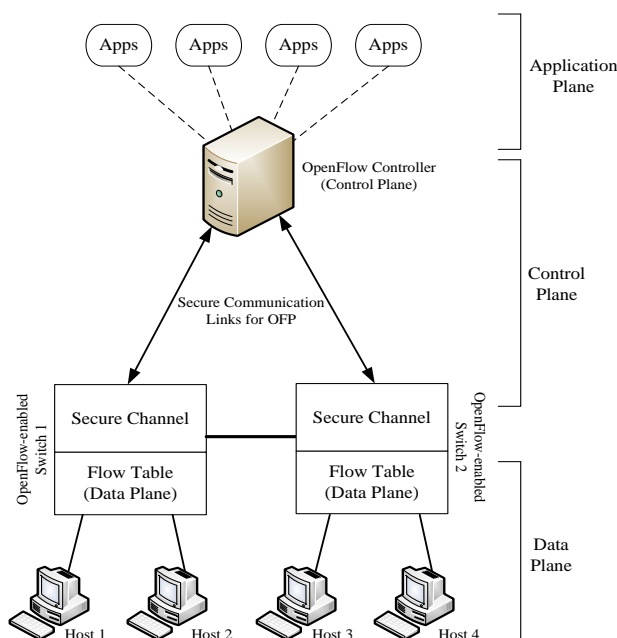


Fig. 3 OpenFlow-based Network Architecture.

OpenFlow-based network architecture is having three layers as discussed previously is distinctly shown in fig. 3. A control plane comprising of NOS is acting as an intermediate layer between data plane and application plane. Again the network

application is defined by network administrator at end user machines. The main components of data plane is OpenFlow-enabled hardware switch, similarly main component of control plane is OpenFlow controller as shown in fig. 3. The switch as shown in figure consists of flow table, group table, meter table, a secure channel, etc. The detail of switch specification is given in [12]. The controller is responsible for controlling all the data flows coming towards the ingress port of an OpenFlow-enabled switch by inserting respective flow entries in a flow table of a switch and the switch is only responsible for forwarding the data packets as directed by an OpenFlow controller. Whenever any data packet is coming towards switch, the switch starts matching the header field of an arrived packet with the flow entries present in its flow tables. If matching is found then the necessary actions are executed on the flows and the data packets are shaped towards its defined destination and if the matching is not found then the switch will forward the arrived packet to the next flow-table in a *pipeline processing* [12]. The process repeats until last flow-table has arrived, if still there is no matching then the switch will either forward the newly arrived packet to the controller or it may drop the packet as directed by controller. The flow diagram and working of multiple flow tables in a pipeline processing is discussed in [13]. Also, an optimal solution for scheduling multiple data flows are discussed in [14].

III. DESIGNING AND PERFORMANCE ANALYSIS OF NETWORK TOPOLOGIES

In this section, a basic network topologies are designed using open source simulation tools. A comparative performance analysis between traditional IP-based network and OpenFlow-based network for the same designed topologies are done. The performance analysis is done by comparing OpenFlow network with traditional network on the basis of round-trip propagation delay between end nodes and maximum obtained throughput. Many open source simulation tools are available for all kind of complex network design. Out of these tools, an efficient and popular tools available are NS (simulator) [15] and Mininet [16]. A traditional network topologies are designed using NS2 simulator script and the OpenFlow network topologies designed using an open source prototype network emulator: A Mininet [17]. The installation guidelines and basic hands-on operation of Mininet is explained in [18].

Mininet simulator supports five built-in network topologies as described in [18]. A default topology is Minimal Topology, this predefined topology consists of one OpenFlow kernel switch connected to two host machines and the OpenFlow reference controller on top of the switch. The three basic network topologies, i.e. *Single Topology*, *Linear Topology* and *Tree Topology*, are discussed in this section. A comparative performance analysis between traditional and OpenFlow-enabled networks for proposed customized network topology using Python script executed in Mininet is also done in this section. A

proposed custom topology using Python script in Mininet is described in detail in [13]. Since, OpenFlow network is having centrally controlled architecture, so the number of central control element is unity, whereas the number of OpenFlow-enabled switches and hosts machines can be connected as many as user requirements. The topological comparison between traditional network and OpenFlow-enabled network is discussed as follows:

A) Network Topologies:

i) *Single Topology:*

A single topology in Mininet is similar to star topology in traditional network. In star topology all the host machines are connected with center hub element, similarly Mininet implements a single topology OpenFlow-enabled network in which a center element is replaced by an OpenFlow-enabled switch, the switch in turn connected with central OpenFlow controller (a control plane). A single topology in traditional network and OpenFlow-enabled network having 16 host machines is shown in fig. 4 and fig. 5 respectively.

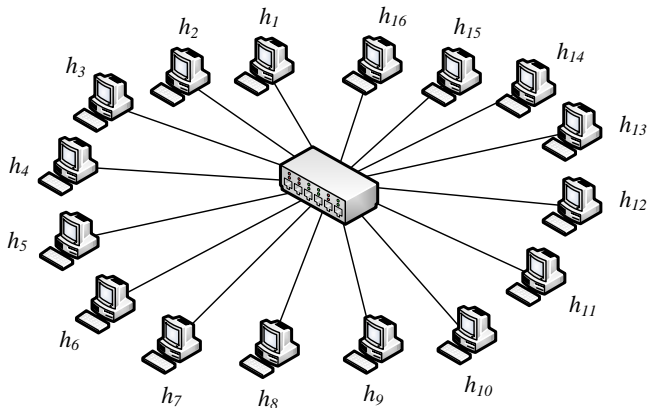


Fig. 4 Single Topology for Traditional Network

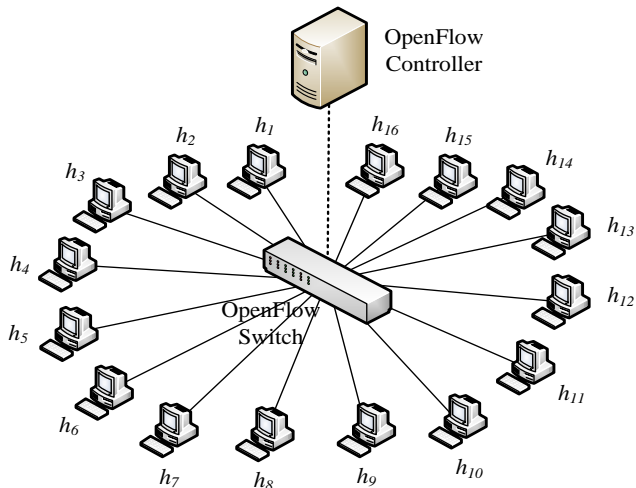


Fig. 5. Single Topology for OpenFlow-enabled Network

As shown in fig. 5, a control plane and data plane is separated from network core device as compared to traditional network design of fig. 4.

ii) *Linear Topology:*

A linear topology in Mininet is similar to Bus Topology in traditional network design with some modifications. In bus topology of traditional network, all the nodes are connected to a single backbone cable called bus with the help of interface connectors. All the workstation can communicate with each other using the common bus. Whereas, in linear topology of Mininet, instead of this backbone cable and interface connectors for all nodes, an OpenFlow-enabled switch and separate cables are used. Moreover, a central OpenFlow controller (a control plane) is present on top of the OpenFlow-enabled switches for controlling data flows. A linear topology for traditional and OpenFlow-based network having 16 host machines is shown in fig. 6 and fig. 7 respectively.

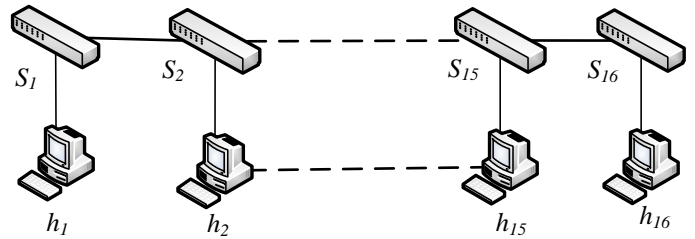


Fig. 6. Linear Topology for Traditional Network

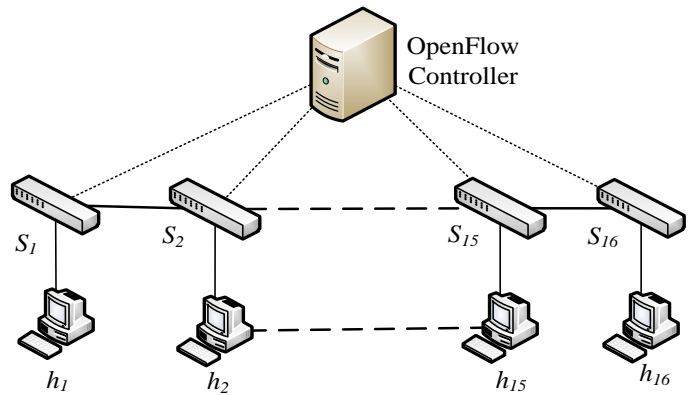


Fig. 7. Linear Topology for OpenFlow-enabled Network

In linear topology of Mininet for traditional network as shown in fig. 6, instead of a common backbone cable and interface connector a separate switch is used for individual host machines and the switches are connected with each other. Moreover, in OpenFlow-enabled network of fig. 7, the traditional switches are replaced by an OpenFlow-enabled switches and these switches in turn gets connected with a centrally controlled OpenFlow controller on a secure link.

iii) *Tree Topology:*

A tree topology in Mininet is slightly complex as compared to single and linear topology as discussed previously. Basically tree topology is based on the amount of depth and the number of fanout of the nodes. That is, the number of switch levels available in the network and the number of output ports available for all nodes. A detail description of tree topology in Mininet is given in [18]. A tree topology for traditional network and OpenFlow-enabled network having 16 host machines is shown in fig. 8 and fig. 9 respectively.

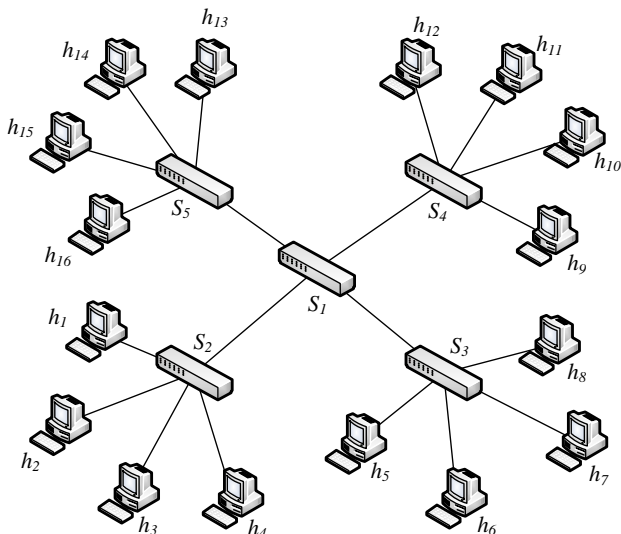


Fig. 8. Tree Topology for Traditional Network

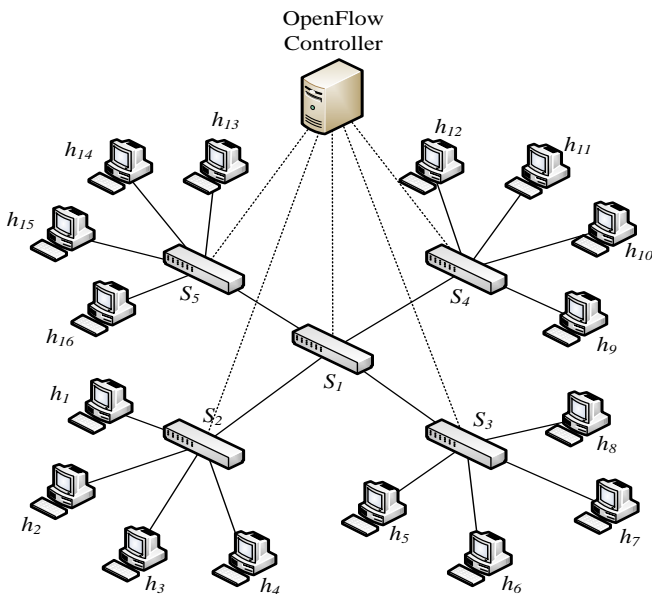


Fig. 9. Tree Topology for OpenFlow-enabled Network

According to previous discussion, a tree topology of Mininet depends on depth and fanout. As shown in tree topology of fig. 8 for traditional network, one can see that the depth is taken as 2 that is 2 levels of switches are present. Level 1 is having switch S_1 and level 2 is having 4 switches S_2, S_3, S_4 and S_5 as shown. After level 2 there is a level of host machines connected with a switch for obvious reasons as shown. Apart from this, tree topology is also having the number of fanouts for the nodes. For this case number of fanout for tree topology is taken as 4, hence the number of output ports for all switches at every level is 4 as shown in the figures.

In tree topology of OpenFlow-enabled network, similar to previous topologies, all the traditional switches are replaced by OpenFlow-enabled switch and the switch in turn connected with OpenFlow controller with a secure link as shown in fig. 9. Finally, it can be said that for designing tree topology having 16 host machines, above configuration for network is required. In a tree topology of Mininet all nodes are connected with one another in a specific hierarchy.

iv) *Custom Network Topology:*

A custom network topology is proposed using Python script supported by Mininet. A backend of Mininet is coded in advance C++ or Python script. Hence, any user defined topology can be implemented using Python script and is executed in Mininet simulator. A proposed topology is having 4 OpenFlow-enabled switch and 32 host machines. A controller resides on top of the switch that controls all the 4 OpenFlow switches. A design and implementation steps for this proposed custom network topology is discussed in [13]. A proposed custom topology for traditional network and OpenFlow-enabled network is shown in fig. 10 and fig. 11 respectively.

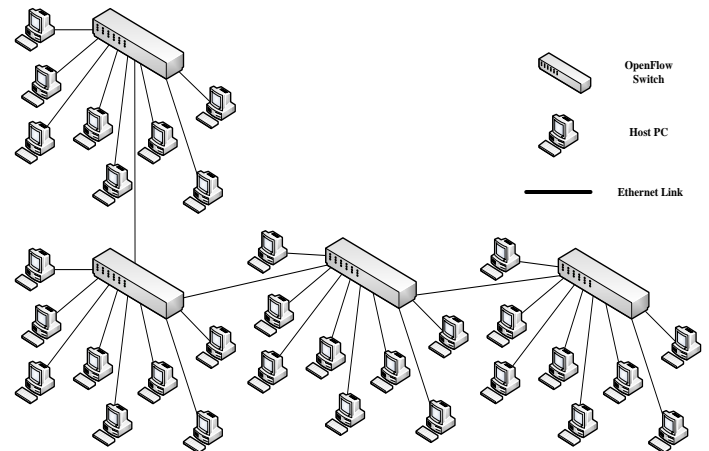


Fig. 10. Mininet Custom Topology for Traditional Network

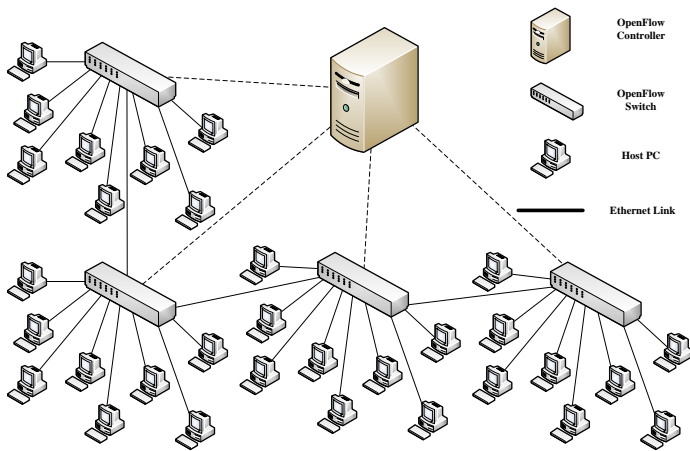


Fig. 11. Mininet Custom Topology for OpenFlow-enabled Network

B) Result Analysis

A comparative performance analysis between traditional network and OpenFlow-enabled network for the above defined network topologies are done by performing network connectivity test between specific nodes in a network and by calculating an average throughput of the network. A simple network connectivity is tested by executing ‘ping’ command that sends ICMP echo request message and wait for its reply to check the IP connectivity between defined nodes and to find round trip propagation delay between nodes. In all the above discussed basic network topologies for traditional network as well as OpenFlow-enabled network, a ping test is performed between node h_1 and node h_{16} . And for proposed custom network topology for both traditional and OpenFlow-enabled network a ping test is performed between end nodes h_1 and h_{32} . A specification of all physical links used for connection between nodes in all network topologies is having 100 Mbps of bandwidth with 1 msec of delay. Based on this physical specifications a ping test is performed between defined nodes of traditional and OpenFlow networks topologies. An obtained round-trip time (rtt) between nodes h_1 and h_{16} for basic topologies and between nodes h_1 and h_{32} for custom network topology is compared and tabulated in Table I and is graphically shown in fig. 12.

TABLE I. ROUND-TRIP TIME BETWEEN NODES

Average round-trip time (rtt) between nodes h_1 and h_{16}		
	Traditional Network	OpenFlow-enabled Network
Single Topology	7.2 ms	4.6 ms
Linear Topology	57.8 ms	41.7 ms
Tree Topology	14.4 ms	8.8 ms
Custom Topology	10.1 ms	8.07 ms

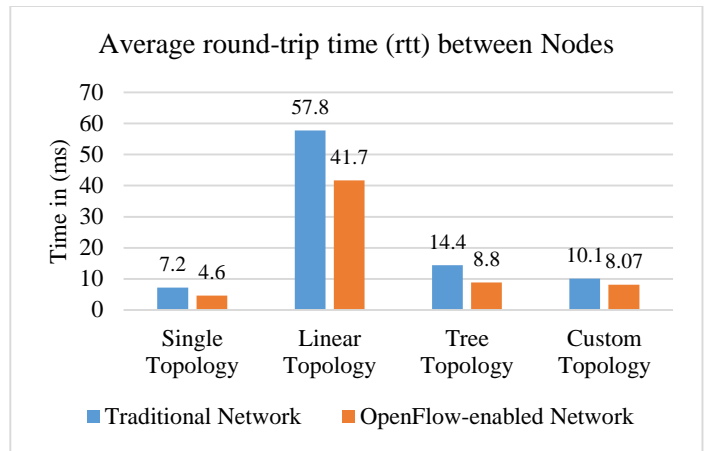


Fig. 12. Average round-trip time between end nodes

A delay between nodes in terms of round-trip time is obtained by performing ICMP query message (echo request and reply) [19] is shown in fig. 12. It is clearly visible that the propagation delay between nodes in OpenFlow network is less as compared to traditional network for any network topology. Next, a comparison between the same traditional network and OpenFlow-enabled network for above discussed topologies is done analyzing obtained throughput of a network. A network throughput is defined as the amount of data transmitted from source node to destination node in a given time period. Throughput is typically measured in bits per second (bps). Analytically it can be defined as the ratio of maximum receiver bandwidth to round-trip time between nodes:

$$\text{Throughput} = \text{maximum receiver bandwidth} / \text{round-trip time}$$

On the basis of available receiver bandwidth and obtained round-trip time (rtt) from previously executed connectivity test, an average throughput is calculated and is shown in fig. 13.

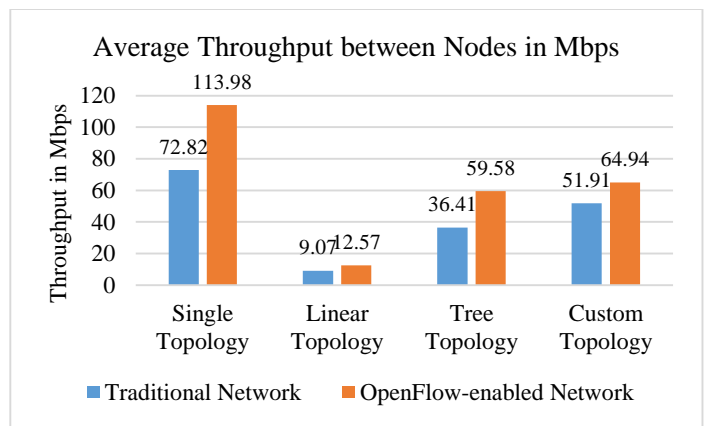


Fig. 13. Average Throughput between Nodes

Here, throughput is calculated between end hosts h_1 and h_{16} in basic Mininet topologies and between h_1 and h_{32} in custom network topology. According to the results obtained as shown in

fig. 13, an average throughput between given nodes is more in OpenFlow-enabled network as compared to traditional network for all topologies. It means that more amount of data transfer take place between nodes in a specific time period for OpenFlow networks then traditional networks.

IV. PROPOSED CAMPUS NETWORK ARCHITECTURE

In this section, a part of the campus network model is proposed by interfacing virtual hosts with real hosts. A comparative performance analysis of traditional network with OpenFlow-enabled network for the same proposed campus network model is also done in later part of this section. Here, a virtual network is created using Mininet simulation tool in a single physical host-only machine such as desktop/laptop. In this virtual network, one of the virtual host is configured in such a manner that it can work as a real host, and this virtual host is then allowed to interface with other real hosts in physical network. A proposed campus network model is shown in fig. 14.

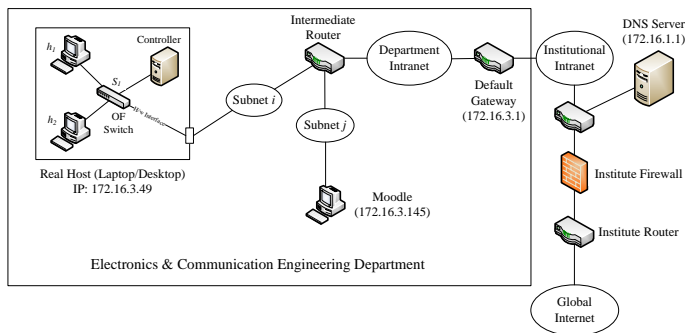


Fig. 14. Proposed Campus Network Architecture

An architecture shown in fig. 14 is a part of the SVNIT-campus network. A virtual network is created using Mininet simulation tool in one of the real host machine located in Electronics and Communication Engineering Department having IP address of 172.16.3.49 as shown. A minimal network topology [17] is created using Mininet command in this real host machine, this topology is having two virtual hosts h_1 and h_2 connected with a single OpenFlow-enabled switch S_1 and the switch in turn connected with OpenFlow controller with separate control plane on top of the OpenFlow network as shown in Figure 14. One of the virtual host, say h_1 , is configured as real host with the same IP address of the available real host machine i.e. 172.16.3.49. A snapshot for configuring virtual host h_1 in Mininet console having Linux environment is shown in fig. 15.

```
*** Starting CLI:
mininet> h1 ifconfig h1-eth0 172.16.3.49/24
mininet> h1 route add default gw 172.16.3.1
mininet> sh echo "nameserver 172.16.1.1" > /etc/resolv.conf
```

Fig. 15. Snapshot of Mininet console configuring virtual host h_1

Besides a real host machine with given IP address, there are other real nodes available in the architecture as shown in fig. 14. A Moodle server node with IP address of 172.16.3.145, a department default gateway node with IP address of 172.16.3.1, an institute DNS server located outside the department but is available inside campus having IP address of 172.16.1.1 as shown.

In the above proposed network architecture, when we only deals with real nodes, then the proposed network itself is an infrastructure traditional network available currently. And when we configure OpenFlow virtual host inside the real host, a part of the proposed network will act as an OpenFlow-enabled network with separate data plane and control plane. Hence, a comparative performance analysis between traditional network and OpenFlow-enabled network for a proposed network architecture is done by performing ping test between nodes and by analyzing round-trip propagation delay between nodes. First ping test in traditional network is performed between real host (IP: 172.16.3.49) and Moodle Server (IP: 172.16.3.145) and for OpenFlow network first ping test is performed between configured virtual host h_1 (IP: 172.16.3.49) and Moodle Server (IP: 172.16.3.145). Similarly, second ping test in traditional and OpenFlow-enabled network is performed between real host and department default gateway (IP: 172.16.3.1) and between virtual host h_1 and the same department default gateway respectively. Finally, a third ping test in traditional network and OpenFlow network is performed between real host and DNS Server (IP: 172.16.1.1) and between virtual host and the same DNS server respectively. A snapshot for all the above three tests performed in traditional as well as OpenFlow-enabled network is shown in below figures (fig. 16 - fig. 21) respectively.

```
idris@ubuntu: ~
idris@ubuntu:~$ ping -c10 172.16.3.145
PING 172.16.3.145 (172.16.3.145) 56(84) bytes of data:
64 bytes from 172.16.3.145: icmp_req=1 ttl=63 time=1.46 ms
64 bytes from 172.16.3.145: icmp_req=2 ttl=63 time=5.07 ms
64 bytes from 172.16.3.145: icmp_req=3 ttl=63 time=5.75 ms
64 bytes from 172.16.3.145: icmp_req=4 ttl=63 time=5.16 ms
64 bytes from 172.16.3.145: icmp_req=5 ttl=63 time=1.62 ms
64 bytes from 172.16.3.145: icmp_req=6 ttl=63 time=2.13 ms
64 bytes from 172.16.3.145: icmp_req=7 ttl=63 time=2.71 ms
64 bytes from 172.16.3.145: icmp_req=8 ttl=63 time=1.45 ms
64 bytes from 172.16.3.145: icmp_req=9 ttl=63 time=1.40 ms
64 bytes from 172.16.3.145: icmp_req=10 ttl=63 time=5.75 ms

--- 172.16.3.145 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9014ms
rtt min/avg/max/mdev = 1.401/3.254/5.756/1.831 ms
idris@ubuntu:~$
```

Fig. 16. Ping test in traditional network between real host and Moodle Server


```

root@ubuntu: /home/idris/Desktop/Paper-3-pics
mininet> h1 ping -c10 172.16.3.145
PING 172.16.3.145 (172.16.3.145) 56(84) bytes of data.
64 bytes from 172.16.3.145: icmp_req=1 ttl=64 time=2.73 ms
64 bytes from 172.16.3.145: icmp_req=2 ttl=64 time=0.602 ms
64 bytes from 172.16.3.145: icmp_req=3 ttl=64 time=0.288 ms
64 bytes from 172.16.3.145: icmp_req=4 ttl=64 time=0.281 ms
64 bytes from 172.16.3.145: icmp_req=5 ttl=64 time=0.283 ms
64 bytes from 172.16.3.145: icmp_req=6 ttl=64 time=0.191 ms
64 bytes from 172.16.3.145: icmp_req=7 ttl=64 time=0.290 ms
64 bytes from 172.16.3.145: icmp_req=8 ttl=64 time=0.269 ms
64 bytes from 172.16.3.145: icmp_req=9 ttl=64 time=0.446 ms
64 bytes from 172.16.3.145: icmp_req=10 ttl=64 time=0.257 ms

--- 172.16.3.145 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9002ms
rtt min/avg/max/mdev = 0.191/0.563/2.731/0.731 ms
mininet>
    
```

Fig. 17. Ping test in OpenFlow-enabled network between virtual host and Moodle Server

```

root@ubuntu: /home/idris/Desktop/Paper-3-pics
mininet> h1 ping -c10 172.16.1.1
PING 172.16.1.1 (172.16.1.1) 56(84) bytes of data.
64 bytes from 172.16.1.1: icmp_req=1 ttl=63 time=1.70 ms
64 bytes from 172.16.1.1: icmp_req=2 ttl=63 time=0.738 ms
64 bytes from 172.16.1.1: icmp_req=3 ttl=63 time=0.305 ms
64 bytes from 172.16.1.1: icmp_req=4 ttl=63 time=0.291 ms
64 bytes from 172.16.1.1: icmp_req=5 ttl=63 time=0.272 ms
64 bytes from 172.16.1.1: icmp_req=6 ttl=63 time=0.417 ms
64 bytes from 172.16.1.1: icmp_req=7 ttl=63 time=0.564 ms
64 bytes from 172.16.1.1: icmp_req=8 ttl=63 time=0.305 ms
64 bytes from 172.16.1.1: icmp_req=9 ttl=63 time=0.312 ms
64 bytes from 172.16.1.1: icmp_req=10 ttl=63 time=0.309 ms

--- 172.16.1.1 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9001ms
rtt min/avg/max/mdev = 0.272/0.521/1.700/0.418 ms
mininet>
    
```

Fig. 21. Ping test in OpenFlow-enabled network between virtual host and DNS Server

```

idris@ubuntu: ~
idris@ubuntu:~$ ping -c10 172.16.3.1
PING 172.16.3.1 (172.16.3.1) 56(84) bytes of data.
64 bytes from 172.16.3.1: icmp_req=1 ttl=255 time=2.34 ms
64 bytes from 172.16.3.1: icmp_req=2 ttl=255 time=68.3 ms
64 bytes from 172.16.3.1: icmp_req=3 ttl=255 time=2.45 ms
64 bytes from 172.16.3.1: icmp_req=4 ttl=255 time=2.28 ms
64 bytes from 172.16.3.1: icmp_req=5 ttl=255 time=2.42 ms
64 bytes from 172.16.3.1: icmp_req=6 ttl=255 time=2.36 ms
64 bytes from 172.16.3.1: icmp_req=7 ttl=255 time=2.43 ms
64 bytes from 172.16.3.1: icmp_req=8 ttl=255 time=8.40 ms
64 bytes from 172.16.3.1: icmp_req=9 ttl=255 time=2.48 ms
64 bytes from 172.16.3.1: icmp_req=10 ttl=255 time=2.23 ms

--- 172.16.3.1 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9014ms
rtt min/avg/max/mdev = 2.231/9.581/68.379/19.681 ms
idris@ubuntu:~$
    
```

Fig. 18. Ping test in traditional network between real host and default gateway

```

root@ubuntu: /home/idris/Desktop/Paper-3-pics
mininet> h1 ping -c10 172.16.3.1
PING 172.16.3.1 (172.16.3.1) 56(84) bytes of data.
64 bytes from 172.16.3.1: icmp_req=1 ttl=255 time=1.19 ms
64 bytes from 172.16.3.1: icmp_req=2 ttl=255 time=1.29 ms
64 bytes from 172.16.3.1: icmp_req=3 ttl=255 time=0.767 ms
64 bytes from 172.16.3.1: icmp_req=4 ttl=255 time=0.984 ms
64 bytes from 172.16.3.1: icmp_req=5 ttl=255 time=0.999 ms
64 bytes from 172.16.3.1: icmp_req=6 ttl=255 time=0.855 ms
64 bytes from 172.16.3.1: icmp_req=7 ttl=255 time=0.777 ms
64 bytes from 172.16.3.1: icmp_req=8 ttl=255 time=0.880 ms
64 bytes from 172.16.3.1: icmp_req=9 ttl=255 time=0.849 ms
64 bytes from 172.16.3.1: icmp_req=10 ttl=255 time=0.794 ms

--- 172.16.3.1 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9003ms
rtt min/avg/max/mdev = 0.767/0.940/1.298/0.172 ms
mininet>
    
```

Fig. 19. Ping test in OpenFlow-enabled network between virtual host and default gateway

```

idris@ubuntu: ~
idris@ubuntu:~$ ping -c10 172.16.1.1
PING 172.16.1.1 (172.16.1.1) 56(84) bytes of data.
64 bytes from 172.16.1.1: icmp_req=1 ttl=63 time=1.51 ms
64 bytes from 172.16.1.1: icmp_req=2 ttl=63 time=1.52 ms
64 bytes from 172.16.1.1: icmp_req=3 ttl=63 time=1.64 ms
64 bytes from 172.16.1.1: icmp_req=4 ttl=63 time=1.46 ms
64 bytes from 172.16.1.1: icmp_req=5 ttl=63 time=1.66 ms
64 bytes from 172.16.1.1: icmp_req=6 ttl=63 time=3.64 ms
64 bytes from 172.16.1.1: icmp_req=7 ttl=63 time=1.65 ms
64 bytes from 172.16.1.1: icmp_req=8 ttl=63 time=2.95 ms
64 bytes from 172.16.1.1: icmp_req=9 ttl=63 time=1.53 ms
64 bytes from 172.16.1.1: icmp_req=10 ttl=63 time=1.52 ms

--- 172.16.1.1 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9016ms
rtt min/avg/max/mdev = 1.463/1.912/3.646/0.716 ms
idris@ubuntu:~$
    
```

Fig. 20. Ping test in traditional network between real host and DNS Server

The above obtained results after execution of ping test is tabulated in Table II and is graphically shown in fig. 22.

TABLE II: ROUND-TRIP TIME BETWEEN NODES

Average round-trip time between nodes in (ms)		
	Traditional Network (IP: 172.16.3.49)	OpenFlow-enabled Network (IP: 172.16.3.49)
Moodle Server (IP: 172.16.3.145)	3.254 ms	0.563 ms
ECED Default Gateway (IP: 172.16.3.1)	9.581 ms	0.940 ms
DNS Server (IP: 172.16.1.1)	1.912 ms	0.521 ms

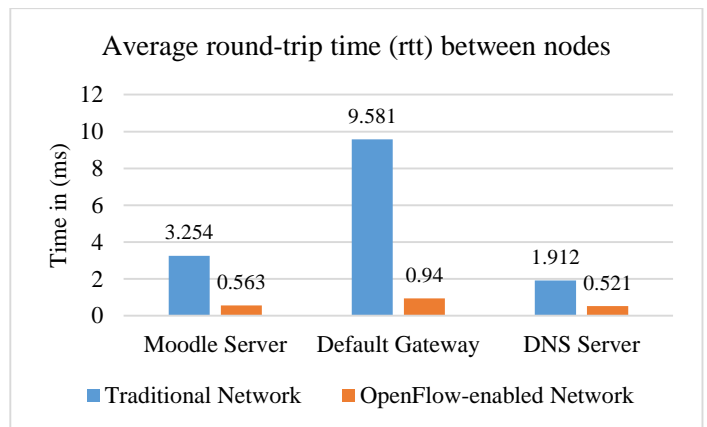


Fig. 22. Average round-trip time between nodes

Hence, from the above results obtained, it is clearly visible that an average time required for OpenFlow-enabled network is less as compared to traditional network. This improvement in round-trip propagation delay will result in increasing speed of data communication and ultimately increases throughput. Thus, by reducing unwanted traffic and overheads of the data packet, a network efficiency gets increases accordingly. The above obtained results will provide us an innovative ideas to deploy a small office or campus network in our institute, so that we can efficiently utilize our network resources.

V. CONCLUSION

A Software-Defined Networking based on OpenFlow technology is the promising technology for future Internet and NGN. Traditional networks are complex in design and are somehow difficult to manage due to the vertical integration of control and data plane on a network core devices. Thus, unlike traditional network architecture, SDN architecture is centrally controlled by separating data plane and control plane from network core devices, this in turn breaks the vertical integration problem of traditional networks.

In this paper a comparative performance analysis of traditional network and OpenFlow-enabled network is done by considering different network topologies. By comparing three basic Mininet network topologies that are single, linear and tree topology, and proposed custom network topology for traditional and OpenFlow-enabled network environment, and on the basis of results obtained it is concluded that the OpenFlow-enabled network is faster due to its centrally control architecture as compared to traditional network. Once a data-flow entry is done in flow table of an OpenFlow-enabled switch by OpenFlow controller for any particular flow, then the switch will keep record for that particular data flow entry for some period of time defined by controller itself. Whenever, successive packets of the same data flow arrives at switch after some time but within the defined time period, then switch will directly forwards the packet to its specific destination on basis of defined flow entry. This will reduce time for route calculation and ultimately a network performance will improve. Moreover, overall throughput of the OpenFlow-enabled network will increase due to fastest data transfer rate.

A virtual node interfacing with real node in a campus network architecture is proposed in this paper. A part of campus network architecture is utilized for a proposed OpenFlow-enabled network in real-time basis. A comparative performance analysis for a proposed network is done between traditional network and OpenFlow-enabled network environment. A ping test is performed between nodes for finding round-trip time (rtt) between defined nodes. On the basis of results obtained it is concluded that OpenFlow-enabled network provides better performance in real-time environment as compared to traditional network. Hence, in concern to future networking scenario, a deployment of OpenFlow-enabled network in small office or campus area is required to be done efficiently so that the networks can become faster, scalable, reliable, and secure, with many more advance features.

REFERENCES:

- [1] T. Benson, A. Akella, and D. Maltz, "Unraveling the Complexity of Network Management," in *Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation*, ser. NSDI'09, Berkeley, CA, USA, 2009, pp. 335-348.
- [2] D. Kreutz, F. M. V. Ramos, P. Verissimo, C. E. Rothenberg, S. Azodolmolky and S. Uhlig, "Software-Defined Networking: A Comprehensive Survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14-76, January 2015.
- [3] S. Paul, J. Pan and R. Jain, "Architectures for the Future Networks and the Next Generation Internet: A Survey," *Computer Communications*, vol. 34, no. 1, pp. 2-42, January 2011, Elsevier.
- [4] A. Gavras, A. Karila, S. Fdida, M. May and M. Potts, "Future Internet Research and Experimentation: The FIRE Initiative," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 3, pp. 89-92, July 2007.
- [5] J. Pan, S. Paul and R. Jain, "A Survey of the Research on Future Internet Architectures," *IEEE Communications Magazine*, vol. 49, no. 7, pp. 26-36, July 2011.
- [6] "Software-Defined Networking (SDN) Definition - Open Networking Foundation," Open Networking Foundation, 1 May 2013. [Online]. Available: <https://www.opennetworking.org/sdn-resources/sdn-definition>.
- [7] M. Jammal, T. Singh, A. Shami, R. Asal and Y. Li, "Software-Defined Networking: State of the Art and Research Challenges," *Computer Networks*, vol. 72, pp. 74-98, October 2014.
- [8] A. Lara, A. Kolasani and B. Ramamurthy, "Network Innovation using OpenFlow: A Survey," *IEEE Communication Surveys and Tutorials*, vol. 16, no. 1, pp. 493-512, First 2014.
- [9] "Home - Open Networking Foundation," 2013. [Online]. Available: <https://www.opennetworking.org/index.php?lang=en>.
- [10] "ONF Overview - Open Networking Foundation," 2013 Open Networking Foundation, [Online]. Available: <https://www.opennetworking.org/about/onf-overview>.
- [11] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker and J. Turner, "OpenFlow: Enabling Innovation in Campus Networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 69-74, April 2008.
- [12] "OpenFlow Switch Specification, Version 1.1.0 (Wire Protocol 0x02)," February 2011. [Online]. Available: <http://archive.openflow.org/documents/openflow-spec-v1.1.0.pdf>.
- [13] I. Z. Bholebawa, R. K. Jha and U. D. Dalal, "Performance Analysis of Proposed OpenFlow-based Network Architecture Using Mininet," *Wireless Personal Communication*, vol. 83, no. 4, pp. 1-18, July 2015, Springer. Online First.
- [14] Y. Liu, Y. Li, Y. Wang and J. Yuan, "Optimal Scheduling for multi-flow update in Software-Defined Networks,"

- Journal of Network and Computer Applications*, vol. 54, no. C, pp. 11-19, August 2015.
- [15] “The Network Simulator – NS-2,” [Online]. Available: <http://www.isi.edu/nsnam/ns/>.
- [16] T. M. Team, “Mininet: An Instant Virtual Network on Your Laptop (or Other PC),” August 2012. [Online]. Available: <http://www.mininet.org>.
- [17] B. Lantz, B. Heller and N. McKeown, “A Network in a Laptop: Rapid Prototyping for Software-Defined Networks,” in *Hotnets-IX Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, New York, NY, USA, October 2010.
- [18] Dr. R. K. Jha, P. Kharga, I. Z. Bholebawa, S. Satyarthi, Anuradha, and S. Kumari, “OpenFlow Technology: A Journey of Simulation Tools,” *International Journal of Computer Network and Information Security (IJCNIS)*, vol. 6, no. 11, pp. 49-55, October 2014.
- [19] Postel, J., “Internet Control Message Protocol”, RFC 792, September 1981.

Reverse Program Analyzed with UML Starting from Object Oriented Relationships

HAMED J. AL-FAWAREH

Software Engineering Department
Zarka University/Jordan

ABSTRACT

In this paper, we provide a reverse-tool for object oriented programs. The tool focuses on the technical side of maintaining object-oriented program and the description of associations graph for representing meaningful diagram between components of object-oriented programs. In software maintenance perspective reverse engineering process extracts information to provide visibility of the object oriented components and relations in the software that are essential for maintainers.

Keywords: Software Maintenance, Reverse Engineering.

1.0 Introduction

The rapid increase in the use of the analysis and design models using UML in software development and the power of new Model have greatly influenced the features of new object oriented software maintenance tools. Program maintenance is an expensive process where an existing program is modified for a variety of reasons, including corrective maintenance, adaptive maintenance, functional enhancement and efficiency improvement. UML models are very popular because UML is a solution of standardization and utilization of design methodologies. Another advantage of UML models is that it provide different diagram for representing different view of system models[5].

Several tools have been developed in order to help the maintainers understand the software system. These tools focus on the maintenance of programming languages like java and C++. The characteristics of object oriented techniques, such as polymorphism, inheritance, encapsulation and dynamic binding are the main reasons for many maintenance problems [1-3]. In particular polymorphism, dynamic binding, sending messages, and overloading pose a new problem for maintainers. These problems include understanding a software system and relationships between an object oriented components [4]. Analysing program relationships is carried out in a maintenance phase to capture different types of relationships among the software component [4]. Object-oriented features give rise to new relations among object-oriented program entities. For examples, association between classes, and method.

The UML has used to notation and graphical representation for the object and message also use to capturing the Message source to destination generally UMLs model is used to design the different types of diagram before the development of any software. UML models are used to source information in software maintenance [5]. Many UML design artifacts have been used in different ways to perform different kinds of maintenance phase. For instance, UML state interaction diagrams (collaboration and sequence diagrams) have been used to test class interactions. Typically, the complexity of an OO system lies in its object interactions, not within class methods which tend to be small and simple [6].

In this paper, we provide a description of association graphs for representing meaningful associations between entities of object-oriented programs. A formal description of the association relations of interest is given before giving a representative illustration of object-oriented program models. A class model represents software components that will be affected due to any particular modification being made to a certain component. The paper also discusses an approach for extracting object-oriented programs through the use of class model. The approach proposes the construction of an automated tool, this tool to extract the association information from the source code. This tool should be used interactively in the maintenance phase to locate the association of a given component of an object-oriented program.

2.0 REVERSE ENGINEERING

Managing and maintaining of software systems require the availability of information knowledge and abstraction level related to the intended objectives and their associated operations. From the software maintenance perspective, reverse engineering is the process of extracting information concerning software product design [7-8]. Information complemented is essential for a developer and maintainer, when a system processes large and huge modules. This information aids humans to understand by representing a complete visualization of the intended objectives.

Moreover, reverse engineering gives facilities that help to control many managerial problems. These problems occur when the developer builds a large software system, which

contains thousand of statements and thousands of methods. For example, a database contains thousands of modules and many thousands of dependency relationship documents and histories. It is beyond the ability of the maintainers to manually search this database to find a specific module or variable, its dependency relationships, and all the activities in these products. Maintainers usually spend more time reading and modifying huge and complex software systems than building them [9].

Thus, the tasks involved in reverse engineering are analysis of a subject system process which is identifying systems components and their interrelationships, in addition to creating representations of the system at higher level of abstraction [9]. The maintainer spends a lot of time to understand all the activities of the problem and to correct it. These involve scanning the source code, reading the documents, and understanding how to make necessary changes. The aim of reverse engineering is to remove ambiguity in the software and understand the software system in different programming languages with the aspect of facilities, enhancements, redesign, and correctness [10].

3.0 Reverse-Tool for Object Oriented Programs,

In this section, we presents a reverse tool for object oriented programs, which will become of an environment for maintaining object oriented system. This reverse-tool focuses on the technical side of maintaining object oriented program, and thus may be viewed as a model of a larger maintenance environment. This larger maintenance environment may be embedded in a software-engineering model, which may be embedded in the corporate model.

The process in the reverse-tool started by a source code modification request arrived. It may be adaptive, perfective, corrective or preventive. To make the change or enhancement in the software system the maintainer needs to understand the source code and the statement that change will occur on it. Also, the maintainers need to know all the relationships and associations between the software components.

The proposed reverse-tool design for understand the software system by giving all the relationships and dependencies that occurs between the software components. These information will feature as relations which will used by the maintainers to understand the certain part of the software system, which is the most important part to modify and change. In order to reduce time, cost and maintainers effort, reverse-tool environment gives the relationships and dependencies captured, so the maintainers will know all the parts effected by the change and enhance part, so this process will reduce time, cost and the maintainers effort.

The proposed reverse-tool assists the maintainers in understanding system constituents and their relationships as well as in associating between object oriented components in object oriented software systems.

The process in the proposed reverse-tool first reads the source code, modifying the software, which requires two steps. First, understanding the source code, which may requires documentation, code reading and execution, second, when the program is modified. The reverse-tool process represents the relationships between object oriented program in three forms class associations with multiplicities, method associations and statement relations.

3.1 Approach taken

The approach taken in this paper includes the use of several associations to support the process of understanding object-oriented program in the maintenance phase. This approach defines three levels of an association graphs which are class, method and statement levels. We propose the construction of an automated reverse-tool supporting the three levels according to user request. The structure of a reverse-tool consists of three steps: the first step is accepts the source code as input file and parses it. The second step is extracts association information from the parse tree and stores all the program elements according to the previously defined and allows them to be presented graphically. Finally represent the association class, method and statements.

3.2 Program Associations Graph

The association graph classifies all the program components and keeps track of the declarations of classes, methods, and variables.

A reverse-tool is directed graph according to the association represented as $G = (V, E)$ where V represents a set of vertices which represent classes, methods and variables, and E represents a set of edges which represents the association between program entities. Each edge represents one of the previous two categories of associations.

An object-oriented program consists of classes and objects, which define the hierarchy of an object-oriented system. Therefore, an object-oriented program contains three levels of components that is class, method and statement level. Each level is integrated with each other in one hierarchy. The class level represents the top level in the association graph where the method and statement level represents the sub-graph level. Likewise the method represents the top level for the variable graph level.

Class Level: This is the highest level. Each vertex represents a class and the graph represents all the associations among the classes and their components. The associations in this level include the three types of class properties that are Class-to-class association through usage, Class-to-class association through inheritance, and Class-class association through causing side effects.

Method Level: This level includes in addition to the vertex of classes, vertex of the method association. This associations includes method associations on another method, and method associations on a variable.

Statement Level: This level captures all types of associations. All the types and their associations are included that are control associations, statement associations on a method, statement dependence on a variable, variable associations on statement, and variable associations on a method.

The visibility of reverse-tool representation for object oriented programs. The association relations among the classes. This represents the class level of the association graph. A class *A* depends on class *B* through the inheritance (*CCDI: Class-Class association through Inheritance*), class *A* depends on class *C* through usage (*CCDU: Class-Class association through Usage*), and class *C* depends on class *B* through side effects (*CCDSE: Class-Class association for causing Side Effects*). The method level of association graph. In this graph, method *m* in class *A* depends on method *r1* in class *C* (*MDM: Method association on another Method*). This is basically method level of association graph. The graph in Figure 1 also shows five types of associations between the class components. All these associations are related to the statement level (i.e. statement level of association graph). These associations are:

1. Association occurring in statement *C.r1.s* in a method *C.r1* of a class *C* is dependent on a method *A.m* in class *A* (*SDM: Statement association on a method*).
2. Association occurring in a statement *B.n2.s* in a method *B.n2* defined in a class *B* is dependent on a variable *C.w1* defined in a class *C* (*SDV Statement Association on a Variable*).
3. Association in a variable *A.v2* defined in a class *A* is dependent on a statement *C.r1.s* in a method *C.r1* of a class *C* (*VDS: Variable Association on Statement*).
4. Association in method *B.n2* defined in a class *B* is dependent on a variable *A.v1* defined in class *A* (*MDV: Method Association on a Variable*).
5. Association in a variable *C.r2.w2* defined in a class *C* is dependent on a method *B.n1* defined in a class *B* (*VDM: Variable Dependent on a Method*).

Class diagram with association representation as shown in Figure 2 will help in understanding the software system. To understand software system means to understand the system components and the relationships between them. In object-oriented programs the associations between various components of the software are the most important to understand the behavior and the structure of the software system. Furthermore, to understand the software system for the purpose of modification and enhancement, the maintainers are required to fully understand the relations that exist within the software. The structure of a software system is determined by the logical and physical organisation of the source code and the relationships that exist within it. A thorough understanding of a software system is possible if the structure and behavior can be explained. The structure and behavior of a system are mutually dependent aspects; the structure of the system permits the software to behave in the desired way and the behavior that is expected from software is the reason as to why the software is structured in a particular way. Reverse-tool is introduced to provide a graphic visualization of object oriented association. This tool also provides the maintainer with an implied methodology for maintenance task.

The graphical used in Figure 1 notation forms the basis of visually representing various kinds of association between components object oriented programs. These associations according to our basic definition will help trace the impact of any proposed modifications, help in understanding object-oriented systems, and perhaps reduce the resources and efforts required for maintenance activities.

3.3 Proposed Maintenance Diagram

The reverse-tool assists in visualizing system constituents and their relationships, and in binding between object oriented program components. In addition, it provides information with the objective to help users to understand various relationships in the object oriented program. The proposed system represent the reverse Java program and visualize it as shown in Figure 1. First step, reads the Java source, and builds up the tokens of the program. Second step accepts the tokens as an input, and produces a parse tree. The third step produces system structures. The system structures represent a component of the software system as *n-tuple*, the Cartesian product of domains d_1, \dots, d_n which is denoted by $d_1 \times d_2 \times \dots \times d_n$ for $n > 1$, and the sets of all tuples (X_1, \dots, X_n) such that, for any $i=1, \dots, n$, $(X_i \in D_i)$. Furthermore, the reverse-tool automatically extracts the information from the software system and transfers these information to the general maintenance diagram with three level class, method and statement as shown in Figure 1.

The system structure given a data structure is defined as a digraph G of an ordered pair $G=(C, R)$ where C denotes a program components and R denotes a component relations, which belong to Cartesian product $D \times D$, where C and R are finite sets. The data structures that are automatically transfer into the parse tree. The user interface accepts questions posed by users and browses the data-based module depending on the structure of the questions posed. The system structure contains relations between all the components of the software system. These relations can be extracted from the source code of the Java language directly.

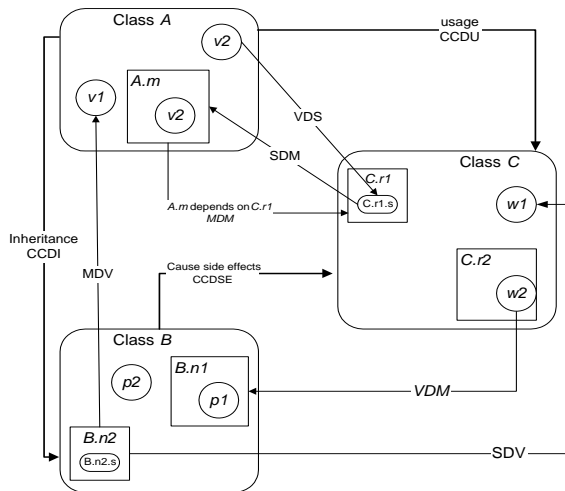


Figure 1: Hierarchical Reverse-tool with the Three Levels

4.0 Case Study

A lot of case studies were performed in order to demonstrate the viability of the proposed tool. Appendix I contains a Java example. It is beyond human ability to search manually a database, which contains thousand of dependency relationships, and find specific modules, their association, and their connectivity with other activities in the life cycle of a product. Furthermore object-oriented technique features a new associations occurs between a program components, that is class, message, methods and variables. Figure 2 shows class diagram for example in appendix I. The class diagram represents the requirement analysis during analysis and design phases. The figure shows the class association relations among the classes. While Figure 3 shows the Proposed maintenance diagram for the same example.

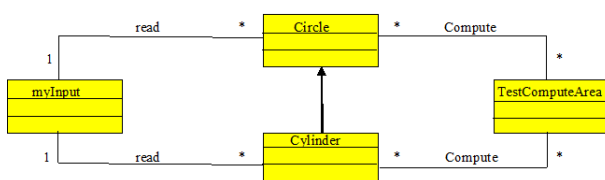


Figure 2: UML class Diagram

The reverse-tool helps the user to search through the system to find any information that specifies the faults locations. Reverse-tool help maintainers to understand a Java source code for a maintenance purposes by given all information about the object oriented system written in Java language. These information include program components that is classes, method, variables and message. Furthermore reverse-tool take into account the associations occurs in object oriented techniques and gives enough information about inheritance, polymorphism, dynamic binding and encapsulations. Figure 2 represent the class diagram as it is in design phase, while Figure 3 represents the class diagram as implemented where obviously, both diagram were almost identical, the difference was an artificial part due to the added information to the original one. The tool also, gives a facility for a maintainer to ask about the object oriented system components not only the declarations or the statements that a specific variables occurs but also the maintainers can ask about the kind of and the statements condition. For example maintainers may ask which object affects a specific variable with a specific condition. The system also, provides report facilities. To illustrate how the user interacts with the system diagram as shown in Figure 3. The diagram contains application associations among classes, attributes and method, which represent the main objectives of the research that is to discover the relationship between “as implemented” and “as designed” architectures.

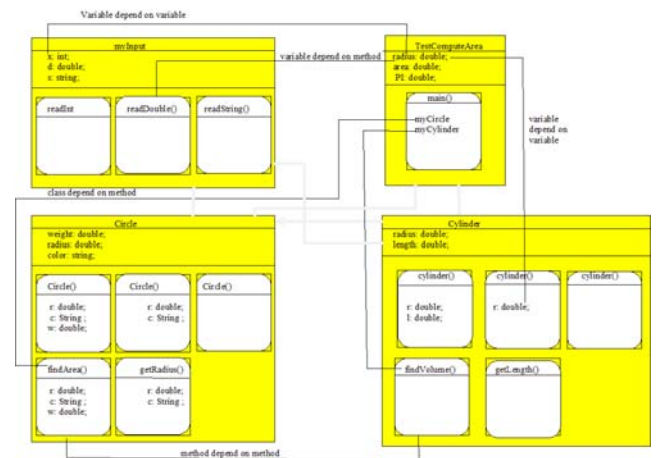


Figure 3: Proposed Maintenance Diagram

5.0 CONCLUSION

We have briefly described the structure and use association graphs in representing meaningful associations between

components of object-oriented programs. An approach to understand the structure and behavior of a given object-oriented program during its maintenance phase. Reverse-tool provides technique to help users, in different way to understand an object oriented system and enhance the software by getting all the information about the classes, methods, variables and their relationships, starting from the tie when users declare it in the program until the end of the class, method and variable life cycle.

REFERENCES

1. Swagatika Dalai, et. al. "Test Case Generation For Concurrent Object-Oriented Systems Using Combinational UML Models" (*IJACSA International Journal of Advanced Computer Science and Applications*, Vol. 3, No.5, 2011).
2. Nabil M. A. Munassar, Govardhan A. "Comparison between Traditional Approach and Object-Oriented Approach in Software Engineering Development" (*IJACSA International Journal of Advanced Computer Science and Applications*, Vol. 2, No. 6, 2011)
3. Ajeet A. Chikkamannur, Shivanand M. Handigund "A Mediocre Approach to Syndicate the Attributes for a Class or Relation", *International Journal of Software Engineering and Its Applications* Vol. 5 No. 4, October, 2011.
4. Ramkrishna Chatterjee, Barbara Ryder, and William Landi. "Complexity of Points-To Analysis of Java in the Presence of Exceptions". *IEEE Transactions on Software Engineering*, Vol. 27, No. 6, June 2001.
5. Wendy Boggs, and Michael Boggs, "UML with Rational Rose", SYBEX Paris, 1999.
6. Grady Booch, James Rumbaugh, and Ivar Jacobson, "The Unified Modeling Language User Guide", Addison Wesley, 1999.
7. Wilde Norman and Huitt Ross, *Maintenance Support for Object-Oriented Programs* IEEE Transaction on Software Engineering, Vol. 18, No. 12, December 1992.
8. Leiter Moises, Meyers Scott and Reiss Steven P. (1992) *Support for Maintaining Object-Oriented Programs* IEEE Transaction on Software.
9. Dan C. Marinescu, et. al. "Understanding the Impact of distribution in object oriented distributed systems using structural dependences" 15th European Conference on Software Maiontenace and Reengineering, 2011, pp. 103-112.
10. Jiun-Liang Chen, Feng-Jian Wang, Yung-Lin Chen "Slicing Object oriented Programs" 20th

Asia-Pacific Software Engineering Conference (APSEC), 2013, pp.395.

Appendix I: Java Example

// Class MyInput

```
public class MyInput
{
    static int x = 10;
    static private StringTokenizer stok;
    int i;
    public static int readInt()
    throw Exception
    {
1         i=0;
2         static private BufferedReader br
           = new BufferedReader( new
           InputStreamReader(System.in), 1);
3         String str = br.readLine();
4         StringTokenizer stok = new
StringTokenizer(str);
5         i = new Integer(stok.nextToken()).intValue();
6         return i;
    }

    public static double readDouble()
    throw Exception
    {
7         double d;
8         d = 0;
9         String str = br.readLine();
10        stok = new StringTokenizer(str);
11        d = new
Double(stok.nextToken()).doubleValue();
        return d;
    }

    public static String readString()
    throw Exception
    {
12        String s = null;
13        String str = br.readLine();
14        stok = new StringTokenizer(str);
15        s = new String(stok.nextToken()).toString();
16        return s;
    }
}

// Class cylinder

public class Cylinder extends Circle {
private double length;
private double radius;
public Cylinder() {
1     length = MyInput.readDouble();
}
}
```



```
public Cylinder(double r, double l) {
2     length = 1;
    }
public Cylinder(double r) {
3     radius = r;
    }

public double getLength() {
4     return length;
    }
public double findVolume() {
5     return findArea()*length;
    }
}

public class TestComputeArea {
static double radius;
static double area;
static double PI = 3.14159;
public static void main(String[] args) {
1     System.out.println("Enter radius");
2     radius = MyInput.readDouble();
3     Circle myCircle = new Circle(radius);
4     System.out.println(myCircle.findArea());
5     Cylinder myCylinder = new Cylinder();
6     radius = MyInput.x;
7     System.out.println(myCylinder.Cylinder(radius));
8     System.out.println(myCylinder.getLength());
9     System.out.print(myCylinder.findVolume());
    }
}

class Circle {
double radius;
String color;
static double weight;
public Circle(double r, String c, double w) {
1     radius = r;
2     color = c;
3     weight = w;
    }
public Circle(double r, String c){
4     radius = r;
5     color = c;
    }
public Circle() {
6     radius = MyInput.readDouble();
7     color = "white";
8     weight = 1.0;
    }

public double getRadius() {
12     return radius;
    }
public double findArea() {
13     color = "Blue";
14     radius = 0;
15     weight = MyInput.x*2;
16     if(radius == 0)
16.1     radius = MyInput.x;
```

```
17     MyInput.x =20;
18     return radius*radius*Math.PI;
    }
}
```



Dr. Hamed Al-Fawareh, received his Bachelor degree in Computer Science from Yarmouk University, Jordan, in 1994. He obtained his M.Sc. in Computer Science from University Putra Malaysia (UPM), Malaysia, in 1998. He Completed his Ph.D. in Software Engineering from University Putra Malaysia (UPM), Malaysia, in 2001. Currently he is an associated professor of software engineering at zarqa university Jordan. He was a dean of the faculty science and information technology at Zarqa University, Jordan from 2010-2013. He is serving as the Secretary General of the Collages of Computing and Information Society (CCIS) at the association of Arab Universities , Editor-in-Chief of the International Arab Journal of Information Technology (IAJIT), and Secretary General of the International Arab Conference of Information Technology (ACIT) from Aug. 2010 to Aug. 2013. His area of interest Software Maintenance, Software Quality, Measurement & Evaluation, Software Reverse Engineering and Reengineering, Software Design Pattern, CARE (Computer Aided Re-Engineering) Tools, Software Testing, and Bioinformatics. He is a member of IEEE and ACM.

Lifetime Optimization in Wireless Sensor Networks Using FDstar-Lite Routing Algorithm

Imad S. Alshawi

College of Computer Science and Information Technology
Basra University
Basra, IRAQ

Ismail O. Alalewi

College of Science
Basra University
Basra, IRAQ

Abstract—Commonly in Wireless Sensor Networks (WSNs), the biggest challenge is to make sensor nodes that are energized by low-cost batteries with limited power run for longest possible time. Thus, energy saving is indispensable concept in WSNs. The method of data routing has a pivotal role in conserving the available energy since remarkable amount of energy is consumed by wireless data transmission. Therefore, energy efficient routing protocols can save battery power and give the network longer lifetime. Using complex protocols to plan data routing efficiently can reduce energy consumption but can produce processing delay. This paper proposes a new routing method called FDstar-Lite which combines Dstar-Lite algorithm with Fuzzy Logic. It is used to find the optimal path from the source node to the destination (sink) and reuse that path in such a way that keeps energy consumption fairly distributed over the nodes of a WSN while reducing the delay of finding the routing path from scratch each time. Interestingly, FDstar-Lite was observed to be more efficient in terms of reducing energy consumption and decreasing end-to-end delay when compared with A-star algorithm, Fuzzy Logic, Dstar-Lite algorithm and Fuzzy A-star. The results also show that, the network lifetime achieved by FDstar-Lite could be increased by nearly 35%, 31%, 13% and 11% more than that obtained by A-star algorithm, Fuzzy Logic, Dstar-Lite algorithm and Fuzzy A-star respectively.

Keywords— Dstar-Lite algorithm, fuzzy logic, network lifetime, routing, wireless sensor network.

I. INTRODUCTION

Wireless sensor network (WSN) is a number of sensor devices (nodes) deployed in an area of interest to monitor one or more physical phenomena such as temperature, light, humidity, movement, etc. Sensors in WSN have the capability of wireless communication in addition to sensing. When has sensed data, a sensor node transmits the data wirelessly to other sensor(s) or directly to data collection unit called base station (sink). Sensor nodes cooperatively perform sensing, data routing and network management tasks. The base station or sink connects the WSN to an existing network infrastructure or to the internet and can be fixed or mobile. One or more sinks can be used in a WSN [1].

The importance of WSNs derived from its wide range usage in various applications including remote monitoring of habitat, battlefield monitoring, and environmental sensing (e.g. temperature, humidity, light, vibration, etc.). In such applications, large number of low-cost sensor nodes are

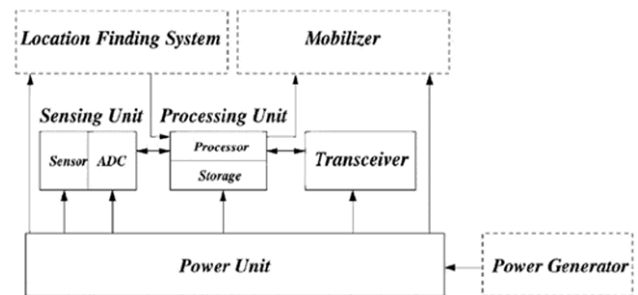


Fig.1 Components of a sensor node

distributed over the monitored area, with nodes organized to form a wireless network, in which each sensor node periodically reports its sensed data to the sink [2] [3].

Generally, sensor nodes in the large-scale data-gathering networks derive their energy from small and inexpensive batteries with limited energy capacity with the expectation of operating for a long period [4]. As shown in Fig. 1, a typical sensor node comprises of sensing, processing, transmission, mobilizing, position finding system and power units. Each sensor node acts upon its mission, the information it currently has, knowledge of its computing, communication, and energy resources.

Having limited transmission range and energy resource of sensor nodes, it is more feasible if sensor nodes send their sensed data through short-distance multiple hops. Hence, the sensor can send its own data or act as a relay to forward data sensed by other sensors in the network. Therefore, energy saving strategies is good solution for power-constrained data-gathering sensor networks. Energy consumption should be given high consideration to maximize the network lifetime [3] [5]. Uneven energy consumption is deep-rooted problem in WSNs characterized by the multi-hop routing and many-to-one traffic pattern. This uneven energy dissipation can significantly reduce network lifetime [4].

Many routing algorithms are share the problem of that they attempt to minimize the total energy consumption in the network at the expense of non-uniform energy drainage in the networks. Such approaches leads network to partition prematurely because some nodes that joins two or more network parts lose their battery energy quicker. According to tradition, the lifetime of a sensor network is over once the

battery power in critical nodes that are shared in many routing paths is depleted. A perfect routing method would enable sensor nodes to share energy consumption slowly and fairly, leading the nodes to die nearly at the same time [2]. Therefore, this paper proposes a new routing method called FDstar-Lite that attempts to investigate the problem of balancing energy consumption, lessening end-to-end delay caused by path planning and maximization of network lifetime for WSNs. FDstar-Lite routing method utilizes Dstar-Lite algorithm alongside with Fuzzy Logic to select an optimal routing path from the source to the destination by preferring the highest remaining battery power and minimum traffic load then reuse the selected path smartly to minimize the delay produced by path planning each time while keeping energy consumption balanced between discovered paths.

The rest of this paper is organized as follows. Related works and concepts (prior arts) of maximizing the lifetime of WSN by routing methods are presented in Part 2. Part 3 describes the problem studied in this paper. In Part 4, the paper gives a brief description of Dstar-Lite algorithm and Fuzzy Logic. The proposed routing method is explained in Part 5. Simulation results are demonstrated in Part 6. Finally, conclusion and discussion are presented in Part 7.

II. RELATED WORK

In conventional routing methods over WSNs, each node selects specific nodes within its transmission diameter as relay for its data under predefined criteria in order to extend network lifetime. Therefore, a good routing scheme in WSNs includes finding the optimal transmission path from the sender, linking one or more relay node(s) toward the destination focusing on increasing network lifetime. Having this idea, the lifetime challenge in WSNs has gained significant attention in the recent past.

The work in [6] proposed to lessen the hop stretch of a routing path (defined as the ratio between the hop distance of the shortest path and that of a given path) in order to reduce the energy consumed by end-to-end transmission. The approaches in [7], [8] have different ways for maximizing the network-lifetime. They tried to distribute the traffic load to the nodes with much residual energy for conserving the availability of the sensors that have less energy. In [6] [7] [8], the works utilized fixed paths for better energy-efficiency; even so, without path variety, those nodes engaged in fixed routing paths may deplete their energy earlier.

ASSORT in [9], offers a joint design that combines the two natural advantages of opportunistic routing, i.e. path diversity and improved transmission reliability alongside with asynchronous sleep-wake scheduling to develop a routing method for extending the network lifetime of a WSN. In [10], a shortest cost path routing algorithm is proposed by Chang and Tassiulas in which link costs that reflect both the communication energy consumption rates and the residual energy levels are taken into account for maximizing network lifetime. The authors in [11] presented a uniform balancing energy routing method where nodes are selected as forwarders for other nodes if their remaining energies are higher than a given threshold in every transmission round, and apportions the energy consumption over many sensors to increase the lifetime of the whole network.

Lu et al. in [12] proposed an Energy-Efficient Multi-path Routing Protocol (EEMRP). It utilizes a load balancing method to assign the traffic over each selected path after searching multiple node-disjoint paths. In this protocol, link costs are calculated based on both the remaining energy level of nodes and the number of hops. The level of load balancing over different multi-paths is evaluated using a fairness index. Furthermore, the reliability of successful paths is sometimes limited since EEMRP only considers data transfer delay. In [13], a novel energy-aware geographic routing protocol (EAGR) in WSNs is proposed. The protocol attempts to reduce the energy consumption of end-to-end transmission. With EAGR, an existing geographic routing protocol is used adaptively to find an anchor list depending on the projection distance of nodes to control the direction of packet forwarding. Each sensor node holding the message makes use of geographic information, the characteristics of energy consumption, and the metric of advanced energy cost to decide forwarding upon and dynamically adjusts its transmission power to the least level required to reach the selected node. The authors of [14] proposed an efficient scheme called data-driven routing protocol to investigate the mobility problems in WSNs with mobile sinks. Data-driven routing protocol attempts to reduce the overheads for the path planning that caused by sink mobility and keep well performing the packet delivery.

On another hand, investing the advantages of Computational Intelligence (CI), a high weight genetic algorithm is used by the protocol presented in [15]. In this protocol, the sensor nodes take into account the data traffic rate to monitor the network congestion. Wang et al. [16] used the Biogeography-Based Optimization (BBO) algorithm to handle the dynamic deployment problem in both static and mobile sensor nodes in WSNs based on a binary detection model. Fuzzy Logic is also utilized in the novel methods presented by [17], [18] and [19] for routing packets in WSNs where each node selects the best node from a set of candidate nodes to form the forwarding paths in order to maximize the lifetime of the sensor networks. Keyur et al. in [20] used A-star algorithm to search optimal path from the source to destination preventing sensor nodes from being engaged in a routing path when their energies is lower than a specific minimum energy level. In [21], Alshawi proposed a routing protocol that incorporates Artificial Bee Colony optimization algorithm (ABC) and Fuzzy Logic, in this protocol, Fuzzy Logic is used alongside with ABC to calculate the optimal path starting from the source node toward the sink considering the traffic load, residual energy and distance to sink to decide which node is the next to be part of the optimal path until the sink is reached. ABC is also exploited by the authors in [22] and [23] to propose a hierarchical clustering scheme for WSNs to maintain least possible energy consumption in the network. Combining Fuzzy Logic with A-start algorithm, Alshawi et al. proposed a novel routing protocol that attempts to lengthen the lifetime of WSNs and balance energy consumption over the network while taking into account the factor of traffic load [24].

III. PROBLEM STATEMENT

Regularly in WSNs applications, numerous sensor nodes are deployed in large areas in dense form. These sensor nodes are equipped with small and inexpensive batteries that cannot be replaced or recharged due to the harsh conditions and inaccessibility of deployment area in most of the applications. Once the limited energy is drained out, nodes will stop working and called “die”. In such a case, the network cannot accomplish its assigned mission or not work in full potential. Therefore, the lifetime is indispensable parameter in WSNs when evaluating the efficiency of routing protocols [3]. The problem of using the same path founded by a routing protocol for next communications over and over as with many routing algorithms in order to achieve battery performance in terms of low transmission delay, the nodes involved in this path will exhaust their energy in fast manner [3] [4] [5]. While these routing algorithms minimize the total energy drainage in the network, they produce uneven energy consumption in the network. Such algorithms cause network partition problem which impairs the usefulness and effectiveness of the whole network [3]. Fig. 2 shows how the network partitioned (a set of nodes may become unreachable) caused by the death of some sensor nodes that are the only connector of a part of the network to the destination.

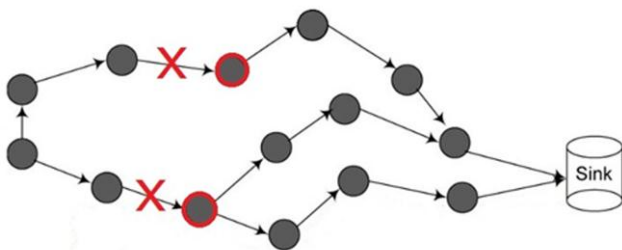


Fig.2 Network partition due to the death of certain nodes

A WSN lifetime is end once the battery power of critical nodes that act as a relay for a part of the network is depleted. The problem then, is to specify a set of routes for each node to forward data through according to some routing parameters (i.e., the routing configuration) that increase network lifetime. Prolonging the lifetime of WSN can be dealt with as optimization problem. In this optimization problem, the variables are routing parameters at nodes. A sensor node can send its own sensed data or act as a relay to forward data of other nodes. In both cases, the node has a role in delivering this data packet to the sink. When the sink is reachable by node’s transmission range, the packet is transmitted directly in single-hop fashion. Otherwise, the packet is sent cooperatively toward the sink by one or more intermediate nodes (hops). In multi-hop transmission, the nodes select one of its neighboring nodes as next hop, that will leads nodes selected as next hop to consume energy in routing the data of other nodes. This case will impact the energy depletion of the entire network and the lifetime as well.

The preceding literatures gives a number of different metrics have been considered to enhance the lifetime of the sensor networks. Some of these are:

a) *Residual Energy (RE)*: To prolong the lifetime of the network, a routing method is needed to consider the factors that have impact on energy consumption. Hence, the most important aspect of routing in WSNs is the energy efficiency. Under this criterion, the focus is on the residual energy (i.e. the current battery power status) of the nodes. A routing method that uses this metric would then attempts to find routes that have the largest total remaining energy from source to sink. In other words, nodes having more remaining energy would participate more than the ones with less power [4] [12] [25]. An example is shown in Fig. 3 where a small sensor network in which, a source node tries to send a packet to a sink node. The numbers inside the nodes refers to the remaining energy (RE) of the nodes. In this example, a routing protocol would choose (1-5-8-11) as the optimal path because it traverses the nodes with largest total remaining energy (i.e. 19).

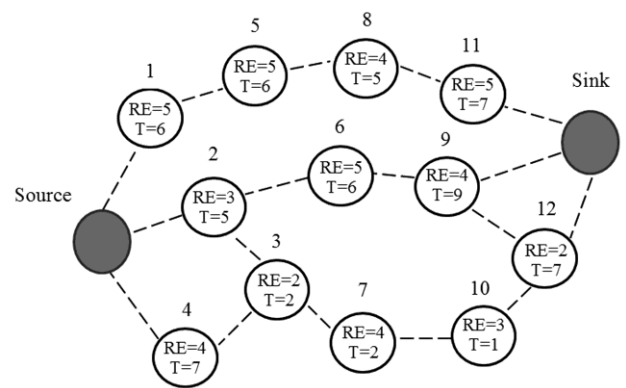


Fig.3 Routing options in a small WSN using different metrics (numbers inside each node indicate the Remaining Energy (RE) and Traffic Load (T) for the corresponding node)

b) *Minimum Hop (MH)*: Minimum hop (or shortest hop) is a well-known criterion used in routing protocols. Under this criterion, the path from the sender (i.e. source) to the sink that traverses the smallest number of intermediate nodes (hops) will be selected as the optimal one. The shortest path will reduce end-to-end delays and resource consumptions by involving the smallest possible number of nodes saving the energy of nodes that are not included in the routing path [12] [19] [25]. In Fig. 3, a routing protocol built on this criterion, could find the path (2-6-9) which comprises the minimum hop (i.e. 3).

c) *Traffic Load (TL)*: The current amount of traffic still present in a node's queue is called traffic load (or intensity) of a node. A data queue overflow problem in the sensor nodes is resulted by high traffic load and could lead to the loss of valuable information. Also, since the limited battery energy of the sensor nodes is quickly depleted when having high traffic load, the lifetime of the whole network would become shorter [4] [19] [25]. Therefore, it's preferred to avoid sending more traffic to the node of high traffic load in order to keep the WSN performing for more time. In Fig. 3, *T* value inside each node represents the traffic load of that node. Considering this criterion, the path (2-3-7-10-12) is the selected one as it totally has the lowest traffic load (i.e. 17).

To extend the network lifetime and reduce the path finding time in each transmission round, this paper proposes a new routing method named FDstar-Lite algorithm. The proposed routing method is used to find the optimal routing path from source toward destination as per the said metrics (Remaining Energy, Minimum hop and Traffic Load) and balancing among them. Also, the proposed method is reusing the path repeatedly for later transmissions until a specific change occur in the cost of any node of that path. Then, another path will be searched starting from the changed node position in order to lengthen the lifetime of the sensor network as much as possible and avoid uneven energy consumption while minifying the delay caused by path finding process.

IV. DSTAR-LITE ALGORITHM AND FUZZY LOGIC

A. Dstar-Lite Algorithm

Dstar-Lite is an incremental heuristic graph search algorithm. An incremental search aims to recalculate only those start distances (that is, distances from the start node to a node) that have changed or have not been calculated before while heuristic search attempts to recalculate only those start distances that are needed for recalculating a shortest path from the start vertex to the goal vertex [26]. The search is done in backward manner (starting from the goal node and ending on start node) evaluating (or expanding) each encountered node based on the sorted priority queue until the start node is reached. The flowchart in Fig. 4 illustrates the general idea of Dstar-Lite algorithm.

Each node has a *Key* value (evaluation function) depends on the two values. These are *g* value, refers to the start distance of a node (node cost) and *rhs* (right-hand side) value which is calculated as the following:

$$rhs(s) = \begin{cases} 0 & s = S_{goal} \\ \min_{u \in succ(s)} (g(u) + c(u, s)) & otherwise \end{cases} \quad (1)$$

where *succ*(*s*) is the set of nodes that are successors to *s* and *c*(*u*, *s*) is the cost of moving from node *u* to node *s*.

The priority queue is sorted based on *Key* values and *Key* of a node (*s*) is a vector of two components calculated as the following:

$$key = [\min(g(s), rhs(s)) + h(s, S_{start}); \min(g(s), rhs(s))] \quad (2)$$

where *h*(*s*, *Sstart*) is an estimate of the distance from node *s* to the start node (*Sstart*).

One advantage of Dstar-Lite algorithm is that it uses knowledge from previous searches to perform later searches much faster than is possible by solving each search task from the beginning. Dstar-Lite algorithm efficiently searches a shortest path from its current node to the goal node by evaluating only those goal distances that have changed (or have not been calculated yet) and it is relevant for recalculating the shortest path. The challenge is to determine these cells efficiently [26], thus we applied a specific mechanism in this regard as explained in the next part.

B. Fuzzy Logic

Since the invention of Fuzzy Logic in 1965 by Lotfy Zadeh, it is widely used by many different applications to enhance calculation processes in such a way that facilitate human knowledge in a specific field by representing that knowledge in a computerized form. A fuzzy control system can be defined as a knowledge-based system, implementing expertise of a human operator or process engineer that can be easily represented through a set of linguistic terms such as “cold”, “normal” or “hot” called fuzzy rules (IF-THEN rules) [27].

In Fuzzy Logic, each object *x* belongs to a set of objects *X* called universe of discourse in a degree. This degree of belongingness is known as membership function. A fuzzy set *A* in *X* is defined by a set of ordered pairs:

$$A = \{(x, \mu_A(x) / x \in X)\}, \quad (3)$$

where the function $\mu_A(x)$ is called membership function of the object *x* in *A* [24]. Fig. 5 shows the fuzzy membership function.

Membership functions offers a mapping of objects to a

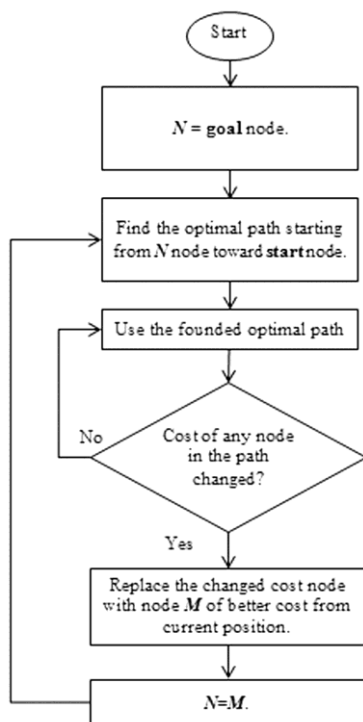


Fig.4 Dstar-Lite (general algorithm)

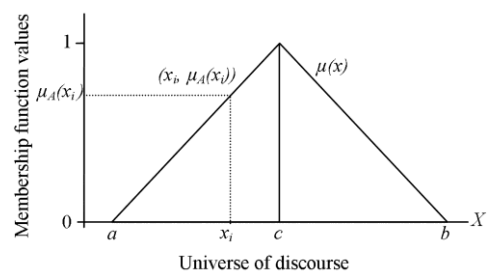


Fig.5 Membership function from the pair (x, $\mu_A(x)$)

continuous membership value in the interval [0...1]. When a membership value is close to the value 1 ($\mu_A(x) \rightarrow 1$), it means that input x is highly belongs to the set A , on the other hand, small membership values ($\mu_A(x) \rightarrow 0$), indicate that set A does not suit input x very well [28]. The dynamic behavior of a fuzzy system is characterized by a set of linguistic fuzzy rules relevant to the knowledge of a human expert. The general form of the linguistic roles is IF antecedent(s) THEN consequent(s), where antecedents and consequents are propositions containing linguistic variables. Antecedents of a fuzzy rule form a group of fuzzy sets through the use of logic operations. Thus, the knowledge base of a fuzzy inference engine is formed by the combination of fuzzy sets and fuzzy rules. Rules are the core of a fuzzy system and may be provided by experts or can be obtained from numerical data. In either case, the rules that we are interested in can be expressed as a collection of IF THEN statements [24].

The fuzzy input space and fuzzy output space are formed by antecedents and consequents of fuzzy rules respectively and are defined by combinations of fuzzy sets. Considering a fuzzy system with p inputs and one output with M rules, the L^{th} rule has the form [28]:

$$R^L : IF x_1 \text{ is } F_1^L \text{ and } \dots \text{ and } x_p \text{ is } F_p^L \text{ THEN } y \text{ is } G^L$$

where $F_1^L \dots F_p^L$ and G^L denote the linguistic variables defined by fuzzy sets and $L=1 \dots M$. Fig. 6 shows the typical structure of a fuzzy system.

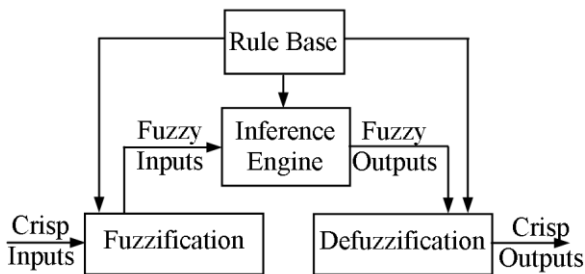


Fig.6 Typical structure of the fuzzy system [27]

The process of turning crisp inputs into fuzzy representations is called fuzzification. This involves application of membership functions such as triangular, trapezoidal, Gaussian etc. The inference engine produces a fuzzy output from the fuzzified inputs by utilizing the rule base. A consequent of the rule and its membership to the output sets is calculated here. Inversely, the defuzzification process involves conversion of the output of a fuzzy rule into crisp outputs using one of defuzzification strategies [24].

V. FDSTAR-LITE ROUTING ALGORITHM IN WSNs

Considering a WSN as a directed-weighted graph, the process of finding an optimal path from the source to the sink in order to forward the sensed data through can be done using a graph search algorithm. Therefore, we proposed FDstar-Lite algorithm that couples Dstar-Lite algorithm with Fuzzy logic. Details of FDstar-Lite implementation in WSNs are given as the following:

A. Implementation of Dstar-Lite Algorithm

This part will describe how Dstar-Lite algorithm is exploited to be applied in WSNs routing. It is used to conserve total network energy for longer time while perform search tasks faster by using knowledge from previous searches to bypass the overheads of initiating path finding process from scratch, therefore, reduce delay of data transmission.

Firstly, we assume that, the sink has knowledge about the current status of each node in terms of battery energy level, location coordinates and traffic load. Thus, the proposed routing method finds the routing path for sensor node that has data to be sent (source node) or (s) toward the sink for the first time as the following:

1. Starting from the sink as current node to be expanded, detect all neighboring nodes that can directly communicates with the sink (i.e., their transmission range can reach the sink). When the source node (s) is detected as sink's neighbor, it could send its collected data directly without any intermediate hop, otherwise, calculate the key values for all detected nodes as the following:
 - a) As assumed, (x, y) coordinates are known for each sensor node in the network, distance (d) of each node (n) to the source node (s) can be calculated as the following:

$$d = \sqrt{(Xs - Xn)^2 + (Ys - Yn)^2} \quad (4)$$

where (Xs, Ys) and (Xn, Yn) are the (x, y) coordinates for nodes n and source node s. The h value for a node can be calculated then by distance normalization to [0...1] using feature scaling method.

$$h = \frac{d - \text{minimum distance}}{\text{maximum distance} - \text{minimum distance}} \quad (5)$$

Knowing the maximum possible distance in the sensor area (i.e., diagonal length) and minimum distance is 0, and then we have:

$$h = \frac{d}{\text{maximum distance}} \quad (6)$$

- b) After that $g(n)$ value represents the remaining energy of node n, that is, battery status of node n.
- c) The $rhs(n)$ value is the current traffics pending in the queue of node n.
- d) Node fitness $NF(n)$ is calculated by applying fuzzy inference system, where $g(n)$ and $rhs(n)$ values are the inputs to that system and $NF(n)$ is the output with a value of range [0..1]. This will be described in the next section of this part.
- e) After h , NF values are calculated, key value is obtained using the following equation:

$$Key = NF + (1/h) \quad (7)$$

2. When a set of nodes detected in the same expansion process, they are successors to the expanded node and substitutes to each other. The pack pointer of each node discovered during the expansion process is set to the expanded node.

3. After each node expansion process, the expanded node will be removed from PQ and the set of detected nodes will be added to the priority queue (PQ) which is then sorted by key values in descending order. The next node to be expanded is the first node in the top of PQ .
4. The process from step1 to step3 is repeated until the source node is detected then the optimal path can be easily traced back by following the back pointers from the source node toward the sink.
5. Path finding process is done when the source node is founded or when the PQ is became empty and the source node is not reached yet. In the latter case the algorithm is failed to find the optimal path due to network partition.

After the optimal path is founded for a specific source node, it will be used for all later data transmissions of that node until change in the cost of any node in the path is detected. Change detection process is done after each time the optimal path is used; it can be explained as the following:

1. Each intermediate node in the optimal path is compared with its best substitute node from step2 in path finding process. In case of the *key* of any intermediate node is smaller than the key of its best substitute node in the list of substitutes, change is said to be detected.
2. Whenever change is detected, the best substitute node becomes the starting point of the next path finding process. Hence the time of path finding from the sink to the changed node position is saved.

B. Implementation of Fuzzy Logic

The function of Fuzzy system used in our proposed routing method is to assign a suitable fitness value $NF(n)$ for each node n depending on $g(n)$ and $rhs(n)$ values of the corresponding node. Fig.7 shows the structure of the fuzzy system used in this routing method.

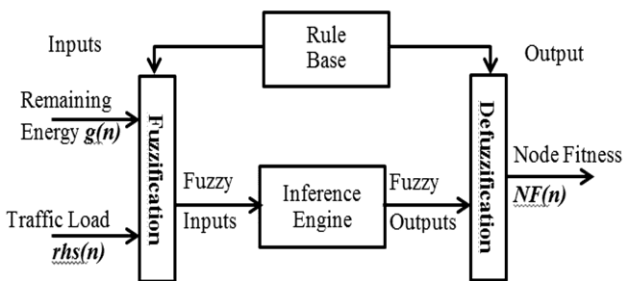


Fig.7 Structure of the fuzzy system for the proposed routing method

There are three membership functions for this fuzzy system, these are $g(n)$ and $rhs(n)$ values as the input membership functions on which fuzzification process is done by mapping these two values to their relevant fuzzy sets. The third membership function is that of $NF(n)$ as the output value of range $[0...1]$. Fig. 8 shows the membership functions of this fuzzy system.

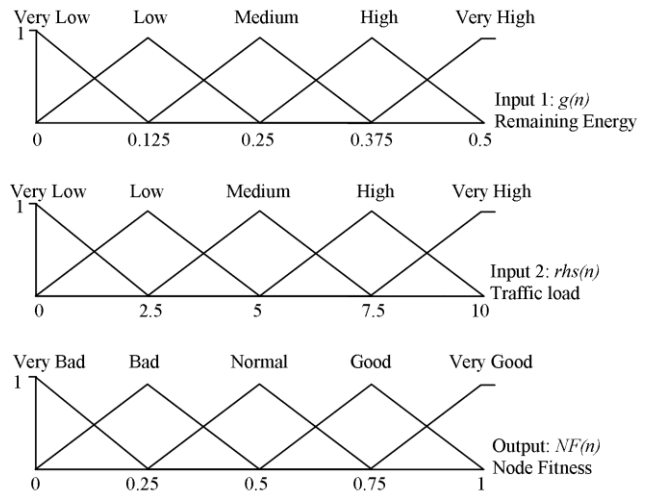


Fig.8 Structure of the fuzzy system for the proposed routing method

Universes of discourse for the three membership functions are $[0...0.5]$, $[0...10]$ and $[0...1]$ for $g(n)$, $rhs(n)$ and $NF(n)$ respectively. After the fuzzification process, the fuzzified inputs will then be passed through the fuzzy inference engine to produce the fuzzy output represented by a linguistic term that belongs to the set of consequents in the If-Then rules illustrated in Table I.

TABLE I. IF THEN RULES

$g(n) \backslash rhs(n)$	Very Low	Low	Medium	High	Very High
Very Low	Bad	Very Bad	Very Bad	Very Bad	Very Bad
Low	Normal	Normal	Bad	Bad	Very Bad
Medium	Good	Good	Good	Normal	Normal
High	Very Good	Very Good	Good	Good	Good
Very High	Very Good	Very Good	Very Good	Very Good	Good

The number of rules in the rule base is $5^2 = 25$ rules as shown in the table above. These rules are used by the fuzzy inference system to determine the output for each fuzzy inputs combination. For example **IF $g(n)$ is High and $rhs(n)$ is Low Then $NF(n)$ is Very Good**. Here, the logical operator of fuzzy implication is AND. These rules are processed in parallel by a fuzzy inference engine. Any rule that fires contributes to the final fuzzy solution space. At the end, the defuzzification finds a single crisp output value from the fuzzy solution space. This value represents the node cost. Practice defuzzification is done using center-of-gravity method [24] given by:

$$NF(n) = \frac{\sum_{i=1}^n U_i * C_i}{\sum_{i=1}^n U_i} \tag{8}$$

where U_i is the output of rule base i and c_i is the center of the output membership function for n rule base number. The flowchart of the proposed routing protocol is shown in Fig. 9.

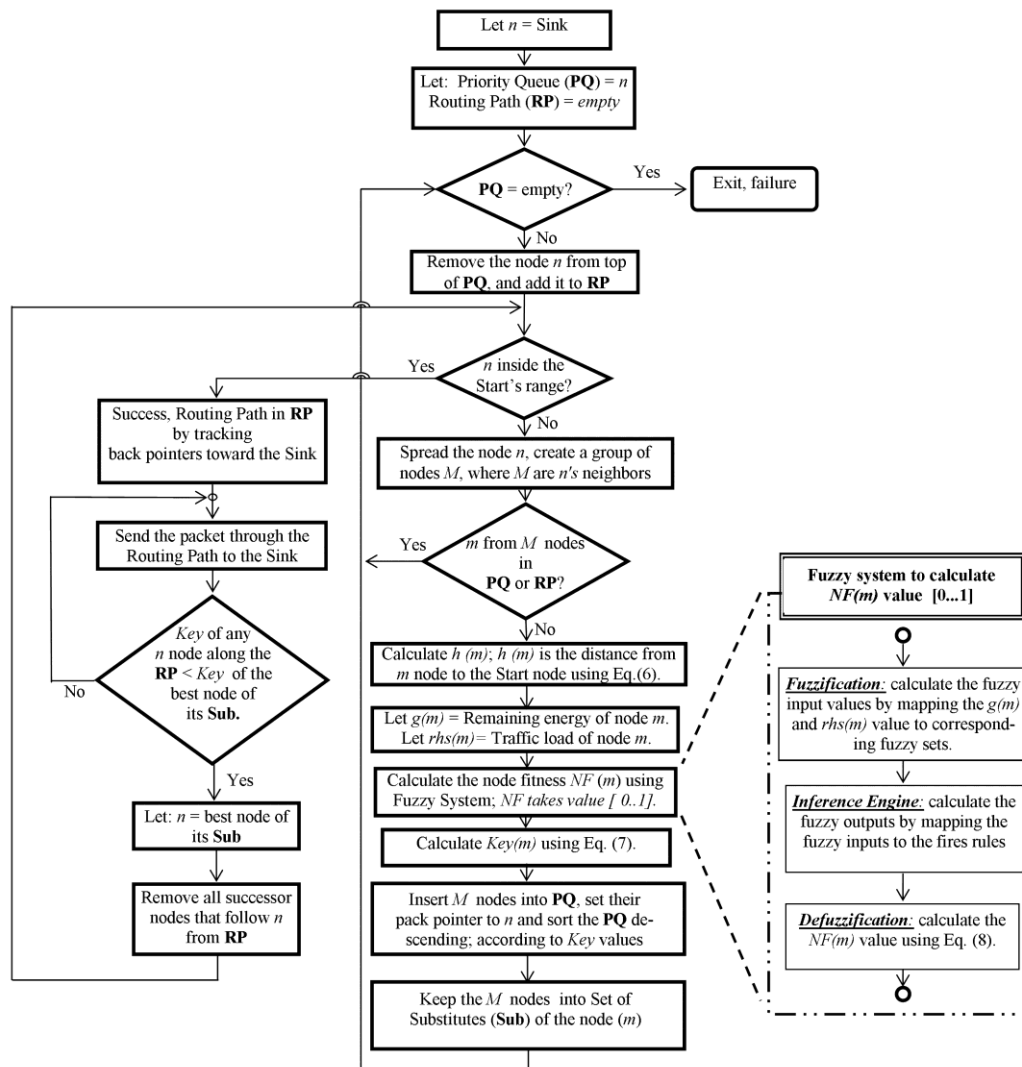


Fig.9 Structure of the fuzzy system for the proposed routing method

VI. SIMULATION RESULTS

To examine the effectiveness of the proposed method considering balancing energy consumption and maximizing network lifetime, simulation results of the proposed are compared with those of A-star algorithm, Fuzzy approach, Fuzzy A-star and Dstar-Lite algorithm. The five methods involve the same routing metrics namely, the remaining energy, the shortest hop, and the traffic load during search of the optimal path from the source node to the sink node.

A-star algorithm is shown to obtain better results than existing maximum lifetime routing algorithms in literatures such as Genetic Algorithm, Warshall algorithm [20] and AODVjr algorithm [29]. The fuzzy approach is also shown to produce better performance over existing maximum lifetime routing algorithms in literatures such as Online-Maximum-Lifetime-heuristic (OML) [17] and Minimum Transmit Energy (MTE) [18]. Moreover, the method proposed in [24] that incorporates both A-star algorithm and Fuzzy approach outperforms both of the A-star algorithm and Fuzzy approach. Dstar-Lite algorithm also performs better than Fuzzy approach and the A-star algorithm. Experimental results obtained under various network scenarios in [17] [18] [20] [24] [29] indicate

that all of the Fuzzy approach, A-star algorithm and Fuzzy A-star give optimal performance in terms of the network lifetime as well as the average energy consumption.

A. Simulation Setup

MATLAB is the simulator software used for this simulation. 100 sensor nodes are deployed in a topographical area that has 100m×100m dimensions. Nodes deployment is done in random manner. The transmission range of each node is limited to (30m). The data sink is located at (90m, 90m). The initial energy capacity is (0.5J) for all sensor nodes in the network. The first order radio model that is frequently used in the area of routing protocol evaluation in WSNs [30] is also used in the proposed method. According to this model, transmission and receiving costs were calculated using the expressions $E_n T(k) = E_{elec} \cdot k + E_{amp} \cdot k \cdot d^2$ and $E_n R(k) = E_{elec} \cdot k$, respectively, where k is the number of bit per packet (packet size), d is the distance from the sender node to the receiver node, E_{elec} is per bit energy consumed in transmitting or receiving circuitry and E_{amp} is energy required per bit per square meter for the amplifier to yield acceptable signal to noise ratio (SNR) respectively. The value assigned in this simulation for E_{elec} is (50nJ/bit) and (100pJ/bit/m²) for E_{amp} .

The traffic load, in each node is represented by a random number generated from [0...10] range of values. Table II gives details of system parameters.

TABLE III. IF THEN RULES

Parameter	Value
Topographical area (<i>meters</i>).	(100m x 100m)
Sink location (<i>meters</i>).	(90, 90)
Number of nodes.	100
Transmission range limit (<i>meters</i>)	30 m
Initial energy of node	0.5 J
E_{elec}	50 nJ/bit
E_{amp}	100 pJ/bit/m ²
Packet size	2k bit
Number of transmitted packets	2×10^4
Maximum node's traffic queue	10

B. Simulation Results

The number of alive nodes for each transmission round using the five different approaches is shown in Fig. 10. It is obvious that the proposed method conserves more alive nodes than A-star algorithm, Fuzzy approach, Fuzzy A-star and Dstar-Lite algorithm after the same number of packets transmitted. When all the 20,000 packets are sent in the area, the network lifetime achieved by the proposed method is nearly 35% more than that obtained by A-star algorithm, 31% more than that obtained by Fuzzy approach, 13% more than that of Dstar-Lite algorithm and 11% more than Fuzzy A-star.

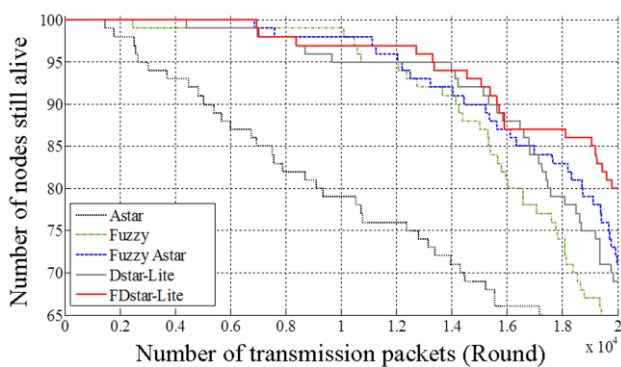


Fig.10 Number of nodes still alive for each round

Moreover in Fig. 10, one can see that our proposed method keeps the number of alive nodes higher than that of the methods compared with.

The number of packet sent when first node died is recorded using the five different approaches as shown in Table III. Remarkably, the first node death occurrence in the proposed method is much later than that of the other four methods. Having Fig. 10 and Table III, one can conclude that, the proposed method is more efficient than A-star algorithm, Fuzzy approach, Dstar-Lite algorithm and Fuzzy A-star in balancing energy consumption and prolonging of network lifetime.

TABLE II. NUMBER OF PACKETS SENT AND FIRST NODE DEATH

Approaches	A-star	Fuzzy	Dstar-Lite	Fuzzy A-star	Proposed method
No. Packet sent causing first node death	1462	2460	4404	6863	6941

The average remaining energy of a WSN decrease with the number of transmission rounds increment. As the number of delivered packets increase, the proposed method results in higher average remaining energy values than A-star algorithm, Fuzzy approach, Dstar-Lite algorithm and Fuzzy A-star. Fig.11 shows that, better energy balance in a WSN is achieved by the proposed method.

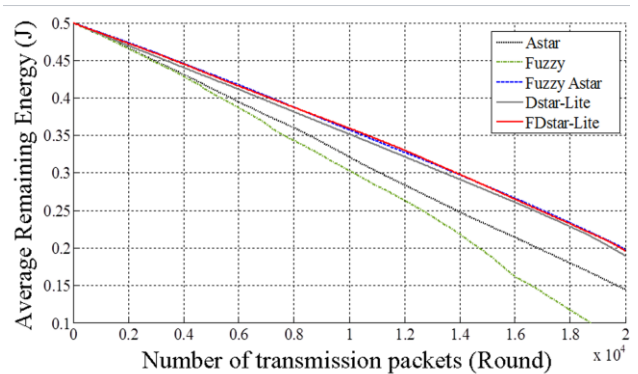


Fig.11 Average remaining energy for each round

The delay caused by data packets transmission is an important parameter for certain applications where sensed data is needed to be collected shortly. The five different approaches are compared as shown in Fig. 12. The proposed method has clearly shortest delay than A-star algorithm, Fuzzy approach, Dstar-Lite algorithm and Fuzzy A-star. This is caused by the nature of Dstar-Lite algorithm that uses knowledge from previous searches instead of starting each path search process from scratch. Shorter delay implicitly indicates energy saving and efficient transmission (especially for secure and important information). Particularly, data packets are sent through different node-disjoint routing paths with multipath routing to eschew network congestion and expand the network lifetime.

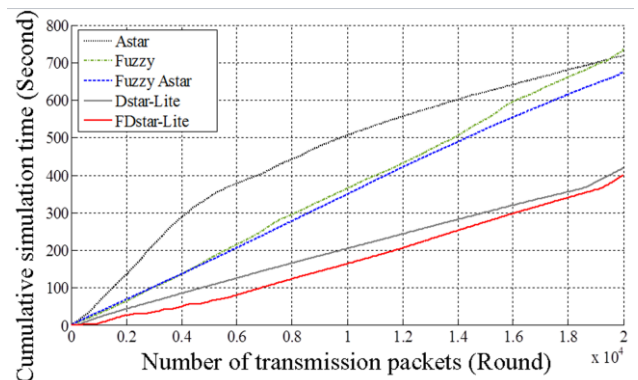


Fig.12 Cumulative simulation time for all packets

Note that above simulations are done with the assumption of that all the nodes are well maintained (i.e. stable with enough power) until the death of node. In real world, there are many factors that lead one or more of the sensors in the critical pathway to perform irregularly. In such case, the parameters of environmental effects and behavior performance noise (fluctuations) into the WSN may be added. As there are too many parameters to be considered, future researches about such topics may be quite interesting and challenging.

VII. CONCLUSION

The limited battery capacity of sensor nodes used in most of WSNs brings the challenge of network lifetime to the top of the list. Therefore, it is important to adopt strategies that efficiently utilize the available energy. Routing path finding methods have a high impact on network lifetime and this is one of the main characteristics of WSNs. Uneven energy drainage is a common problem in a WSN. To achieve efficient data transmission through routing path that is selected to be an optimal path to maximize the overall lifetime of the network with reducing the delay caused by path finding process, we proposed FDstar-Lite, a new method adopting a combination of Dstar-Lite algorithm and Fuzzy Logic. The new method is capable of finding an optimal routing path to be used in data transmission from the source node toward the sink involving intermediate node(s) by preferring nodes with the highest residual energy, minimum hops incorporated and lowest pending traffic. Comparing the proposed method with other four methods, the results demonstrate that the performance of the proposed method, under the same criteria, is much better than that of the four methods in terms of network lifetime and transmission delay.

REFERENCES

- [1] J. N. Al-Karaki; A. E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey," *IEEE Wireless Commun.*, vol. 11, no.6, pp. 6-28, Dec., 2004.
- [2] R. V. Kulkarni., A. Forster., and G. K. Venayagamoorth, "Computational Intelligence in Wireless Sensor Networks: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 13, no. 1, pp. 68-96, Feb., 2011.
- [3] Cunqing Hua, and T. P. Yum, "Optimal Routing and Data Aggregation for Maximizing Lifetime of Wireless Sensor Networks," *IEEE/ACM Trans Networking*, vol.16, no.4, pp. 892 – 903, Aug., 2008.
- [4] Haibo Zhang, and H., Shen, "Balancing Energy Consumption to Maximize Network Lifetime in Data-Gathering Sensor Networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no.10, pp. 1526 - 1539, Oct., 2009.
- [5] H. R. Karkvandi, E. Pecht, and O. Yadid, "Effective Lifetime-Aware Routing in Wireless Sensor Networks," *IEEE Sensors J.*, vol.11, no.12, pp. 3359-3367, Dec., 2011.
- [6] M. J. Tsai, H. Y. Yang, and W. Q. Huang, "Axis-Based Virtual Coordinate Assignment Protocol and Delivery-Guaranteed Routing Protocol in Wireless Sensor Networks," in *IEEE INFCOM*, May., 2007, pp. 2234-2242.
- [7] J. Park, and S. Sahni, "An Online Heuristic for Maximum Lifetime Routing in Wireless Sensor Networks," *IEEE Trans. Comput.*, vol. 55, no.8, pp.1048- 1056, Aug., 2006.
- [8] C. Wu, R. Yuan, and H. Zhou, "A novel Load Balanced and Lifetime Maximization Routing Protocol in Wireless Sensor Networks," *67th IEEE Veh. Tech. Conf.*, 2008, pp. 113-117.
- [9] Hsu, Chih-Cheng, et al. "Joint Design of Asynchronous Sleep-Wake Scheduling and Opportunistic Routing in Wireless Sensor Networks." *Computers, IEEE Transactions on* 63.7 (2014): 1840-1846.
- [10] J. H. Chang, and L. Tassiulas, "Maximum Lifetime Routing in Wireless Sensor Networks," *IEEE/ACM Trans Networking*, vol. 12, no. 4, pp. 609- 619, Aug., 2004.
- [11] O. Zytoune, M. El-aroussi, and D. Aboutajdine, "A Uniform Balancing Energy Routing Protocol for Wireless Sensor Networks," *Wireless Pers. Commun.*, vol. 55, no. 2, pp. 147-161, Oct., 2010.
- [12] Y. M. Lu, and V. W. S. Wong, "An Energy-Efficient Multipath Routing Protocol for Wireless Sensor Networks," *64th IEEE Veh. Tech. Conf.*, pp. 1-5, 2006.
- [13] Haojun, H.; Guangmin, H.; and Fucai, Y.: Energy-aware geographic routing in wireless sensor networks with anchor nodes. *International Journal Communication Systems*, 26(1), 100-113, (2013).
- [14] Shi, L.; Zhang, B.; Mouftah, H. T.; and Ma, J.: DDRP: an efficient data-driven routing protocol for wireless sensor networks with mobile sinks. *International Journal Communication Systems*, DOI: 10.1002/dac.2315, (2012).
- [15] C. Park, and I. Jung, "Traffic-Aware Routing Protocol for Wireless Sensor Networks," in *IEEE ICISA*, Apr., 2010, pp. 1-8.
- [16] Wang, G.; Guo, L.; Duan, H.; Liu, L.; and Wang, H.: Dynamic deployment of wireless sensor networks by biogeography based optimization algorithm. *Journal of Sensor and Actuator Networks*, 1(2), 86-96, 2012.
- [17] M. R. Minhas, , S. Gopalakrishnan, V. C. M. Leung, "An Online Multipath Routing Algorithm for Maximizing Lifetime in Wireless Sensor Networks," in *Proc. IEEE ITNG*, Apr., 2009, pp. 581–586.
- [18] M. A. Azim, and A. Jamalipour, "Performance Evaluation of Optimized Forwarding Strategy for Flat Sensor Networks," *IEEE Global Telecommun. Conf.*, 2007, pp. 710 – 714.
- [19] S. Y. Chiang, and J. L. Wang, "Routing Analysis Using Fuzzy Logic Systems in Wireless Sensor Networks," *Lecture Notes in Computer Science*, vol. 5178, ch. 120, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 966-973.
- [20] K. M. Rana, and M. A. Zaveri, "ASEER: A Routing Method to Extend Life of Two-Tiered Wireless Sensor Network," *Int. J. of Advanced Smart Sensor Netw. Syst.*, vol. 11, no. 2, pp.1-16, Oct., 2011.
- [21] Imad S. Alshawi, "Balancing Energy Consumption in Wireless Sensor Networks Using Fuzzy Artificial Bee Colony Routing Protocol," *International Journal of Management & Information Technology*, vol.7, no.2, pp. 1018-1032, Nov. 2013.
- [22] Karaboga, D.; Okdem, S.; and Ozturk, C.: Cluster based wireless sensor network routing using artificial bee colony algorithm. *Wireless Networks*, 18(7), 847-860, (2012).
- [23] Okdem, S.; Karboga, D.; and Ozturk, C.: An application of wireless sensor network routing based on artificial bee colony algorithm. *IEEE Congress of Evolutionary Computation*, 326-330, (2011).
- [24] I. S. Alshawi, L. Yan, W. Pan, B. Luo. Lifetime enhancement in wireless sensor networks using fuzzy approach and A-star algorithm. *IEEE Sensors J.* 2012, 12(10): 3010-3018. (SCI: WOS: 000307901200011, EI: 20123515374241)
- [25] W. Dargie, and C. Poellabauer, "Network Layer," in *Fundamental of Wireless Sensor Networks Theory and Practice*, 1st. Chichester, UK: John Wiley & Sons, Ltd, Jul., 2010, pp. 163-204.
- [26] Sven Koenig, Maxim Likhachev, "Fast Replanning for Navigation in Unknown Terrain," *IEEE TRANS. Robotics*, vol. 21, no. 3, pp.354-363, Jun 2005.
- [27] L. A. Zadeh. Soft computing and fuzzy logic. *IEEE Software*, 1994, 11(6): 48-56.
- [28] K.-Y. Cai and L. Zhang. Fuzzy reasoning as a control problem. *IEEE Trans. Fuzzy Syst.*, 2008, 16(3): 600-614.
- [29] X. H. Li, S. H. Hong, and K. L. Fang, "WSNHA-GAHR: a greedy and A* heuristic routing algorithm for wireless sensor networks in home automation," *IET Commun.*, vol. 5, no. 13, pp. 1797-1805, Sep., 2011.
- [30] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, 2000, vol. 2, pp. 1-10.



Dr. Imad S. Alshawi (M'12) was born in Basra, Iraq, in 1976. He received the B.Sc. and M.Sc. degrees in computer science from the College of Science, Basra University, Basra, Iraq, in 2001 and 2003, respectively. He received the Ph.D. degree in information and communication system from Southwest Jiaotong University, Chengdu, China. He is currently an Assistant Professor with College of Computer Science and Information Technology,

Basra University. He is a frequent referee for more than 10 journals. He is the author or co-author of more than 30 papers published in prestigious journals and conference proceedings. His current research interests include Wireless Sensor Networks, Artificial Intelligent, Network Management and Information Security.

Dr. Alshawi is a member of the IEEE, the IEEE Cloud Computing Community and the IEEE Computer Society Technical Committee on Computer Communications.



Ismaiel O. Alalewi was born in 1991. He is an M.Sc. student in Computer Science Department, College of Science, University of Basra, located in his hometown of Basra, Iraq. He received the B.Sc. degree from Computer Science Department in Shatt Al-Arab University College, Basra, Iraq, in 2012. He is recently interested in Wireless Sensor Networks.

An Algorithm for Signature Recognition Based on Image Processing and Neural Networks

Ramin Dehgani, Ali Habiboghli

Abstract— Characteristics related to people signature has been extracted in this paper. Extracted Specialty vector under neural network has been used for education. After teaching network, signatures have been evaluated by educated network to recognize real signature from unreal one. Comparing the results shows that the efficiency of this method is better than the other methods.

Index Terms— signature recognition, neural networks, image processing.

I. INTRODUCTION

Handwritten signatures are the most common criterion for recognizing and confirming identity of people. In official system, most sectors especially in commercial domain such as banks, institutes and organizations are the applicants of methods for determining identity of people. Signature is one of the biometric characteristics of people by which person can confirm and approve a document. Hence, recognizing this fact that the signature on the document belongs to which person is so important. Signature recognition method and recognizing is real or not is a popular technique among the users. The advantage of using signature recognition method is that most of new portable computers apply handwritten inputs and don't need for innovation and inventing new hardware systems to collect data. Overall, signature recognition is one of the safest methods of identity recognition and it is accepted in commercial activities [1].

Signature recognition methods have been classified into two main groups: Static and dynamic.

Static method considers signature as a two-dimensional image that doesn't have any information related to time. Therefore, static properties of signature that are not variable by time, are used to recognize signature. In dynamic method of signature recognition, pen movement while writing is

considered and a particular tool called digitalizing paper and a sensitive pen to pressure are applied. In other words, dynamic method uses the dynamic characteristics of signing process. This method involves extracting some characteristics of recorded information in the signing procedure and comparing them with the characteristics of referred and denoted signature [2, 3, 11]. How design a biometric system to recognize a person is an important matter. Based on situation, these systems can be used for recognition and confirming people. Recognition system tries to survey entered biometric data validity regarding the existed ones in database. Some of the identity recognition methods are pointed here:

Token is usually something that you carry with yourself and it can be regarded as your identity document, such as intelligent cards, magnetic cards, key, passport, ID card and so on. All these things have some faults like: being lost, not being with person, being old and being impersonate or counterfeit.

Second type of recognition systems are called knowledge that are things you remember such as password and pin code. These systems have also some faults like forgetting and being transpire.

Third class contains biometric based systems. These systems use human's physiologic and behavioral characteristics for recognition. This method doesn't have the faults of previous classes and has significantly increased safety and accuracy. Measurement and statistical analysis of biologic data is called biometric. Biometric refers to a technology for measuring and analyzing properties of body by using particular attributes (physiological or behavioral attributes) for personal identity recognition.

All the biometric systems have parameters to introduce properties and abilities of systems such as: False Acceptance Rate: This parameter determines the possibility of accepting real user from unreal one. This parameter should be as small as possible.

False Rejection Rate: This scale demonstrates how much person is not accepted by mistake (very high sensitivity). This parameter also should be small enough.

Equal Error Rate: decreasing rate of mistake acceptance increases inevitably rate of mistake unacceptance. The point in which rate of mistake acceptance equals with the rate of mistake unacceptance is the point of Equal error rate. The less amount of this parameter shows that the system has a better sensitivity and an appropriate balance.

Enrollment Incorrect: It is the possible error that may be occur while sampling to record in database to distinguish correctly [4].

This paper submitted 25 February 2016, Revision date is 28 March 2016, acceptance date is 20 March 2016. "This work was supported in part by the Iran, Department of Computer Science and Engineering Departments in Islamic Azad University".

Ali Habiboghli is with Computer Science and Engineering Department, Islamic Azad University, Khoy Branch, Khoy, Iran. Tel. +984436430001-6

Ramin Dehgani is master of science student in Computer Science and Engineering Department, Islamic Azad University, Khoy Branch, Khoy, Iran. Tel. +984436430001-6

The first efforts in simulation by a logic model were done by Mackluk and Walter Pitz which are the main block of constructing most artificial neural networks. This model presents some hypothesis about neurons performance. Performance of this model is based on collecting and sum of inputs and creating output. If the sum of inputs is more than threshold, neuron will be motivated. The result of this model was doing simple functions such as OR and AND.

Not only Neurophysiologists but also psychologists and engineers have influenced the progress of neural networks. In 1958, Perceptron network was introduced by Rouzenblat. This network was similar to the previous modeled units. Perceptron has 3 layers with a middle layer known as a connecting layer. This system can learn to impose proper random output to a given input.

Obtained improvements in 1970 and 1980 were important to pay much attention to neural networks. Some other factors have roles in this case such as wide conferences and books presented to people in different fields. Today, ANN technology has faced with significant changes [5].

The followings are some static methods for signature recognition including: Two-dimensional transformations, directional information histogram, curviness consideration, creating horizontal and vertical image, following signature and finding the location of special places in signature building.

One of the pioneers in this field in 1980s is Ammar that used the idea of statistical consideration of dark points to recognize pseudo dynamic characteristics. That is, there is a positive relationship between darkness level and the amount of grayness with pressing the pen that is one individual attribute. Some Dynamic recognition methods are as follow: Possible classifiers, variation time, neural networks (ANN), and Markov hidden models, signal correlation methods, hierarchical methods, and Euclidean distances and so on[6,9,10].

In this paper we has been extracted Characteristics related to people signature. Extracted Specialty vector under neural network has been used for learning. After teaching network, signatures have been evaluated by educated network to recognize real signature from unreal one. Comparing of the results shows that the efficiency of proposed method is better than the other methods.

This paper is organized following: in the next section artificial neural networks have been considered first in section 3 proposed method has been presented. Discussions include the results of proposed algorithm was explained in section 4. Finally, Conclusion and suggestions are the next sections.

II. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network is a system for processing data. It has used the idea of human brain. Data process is done by a lot of microprocessors that are connected as a network. They act parallel to each other to solve a problem.

Neural Network includes the elements of building the

layers and weights. Network behavior is dependent to the relationships between the members. Overall, there are three kinds of neurons layer in Neural Network: Input layer: receiving the raw information of network; Hidden layers: Performance of these layers is determined by the inputs and the related weight between them and hidden layers.

The weights between entrance and hidden layers determines when should activate a hidden unit. Output layer: Performance of output unit depends on the activity of hidden unit and the related weight of hidden unit and output. Figure 1 shows a section of one neuron layer in neural network.

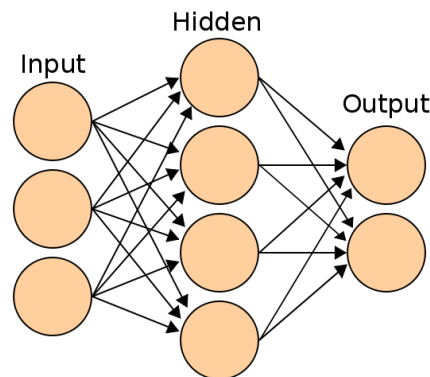


Figure 1. a section of one neuron layer in neural network

Neural networks are divided into 4 groups based on teaching method: Fixed weight: There is no learning and weights amounts are not updated.

Learning without supervisor or headman: weights are only corrected by inputs and there is not any appropriate output to correct them by comparing network output and determining error of weights.

Learning with supervisor or headman: appropriate outputs are shown to network according to the input patterns. Changing weights continues till the output differences of network are acceptable for learning patterns. Reinforcement learning: quality of the system performance is improved step by step in proportion to time [10].

III. PROPOSED METHOD

In this paper, a new signature recognition method is proposed that first by proposing preprocess methods on the images of signatures, an image without noise is created. Then by imposing other methods, we try to obtain the main characteristics of that signature. In the next step, we try to educate neural network by the use of neural network and extracted attribute curve. Finally, the efficiency of proposed method is evaluated. Flowchart of performed procedures is shown in figure 2.

A. Preprocess

As it mentioned earlier, our database is SVC 2004. First we make existed signatures (real or unreal) in this database ready for extracting attributes by preprocess. it can be explained that

preprocess is the process in lower levels. The main objectives of this step is promoting image and removing unnecessary indexes from image (removing noise).

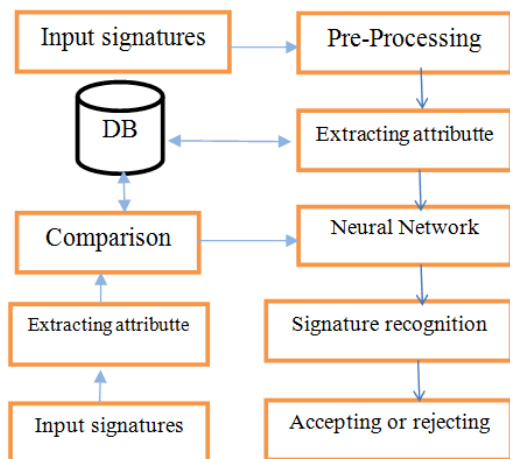


Figure2. Flowchart of Proposed Method

This step has this characteristic that both input and output is the image. The mean filter with mask 3*3 is used for this purpose. The image of mean filters is the simplest calvative filters. A mean filter $m \times n$ is a mask with a number of 1 for each elements (m, n). In other words, the amount of each pixel in output image equals with the mean of light intensity in current pixel and the vicinity pixels. For calming the image a mean filter of 3*3 is used that each element of this filter equals with 9/1. An example of $M \times N$ mask is demonstrated in figure 3.

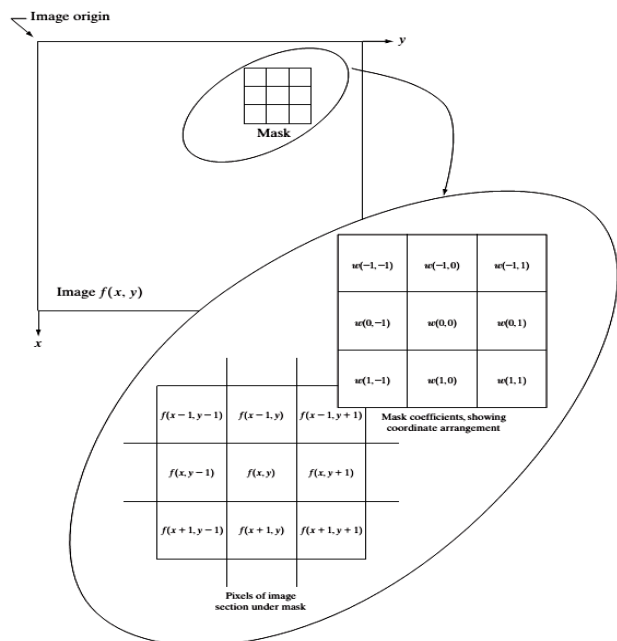


Figure 3. An example of $M \times N$ mask

B. Extracting the attributes

To determine the identity and creator of an image from its pattern, we have to extract some general or specific

characteristics out of image that is called attribution extraction. For example, in signature recognition by image processing, some attributes (like lines inclination) are extracted out of scanned image of signature to recognize the person who has signed it. The main purpose of attribute extraction is to make the raw data ready and usable for the next statistical processes. Different methods of extracting attribute may accomplish one of the following task or more according to the philosophy behind them:

Eliminating data noise, Separating independent elements of data, decreasing the dimension for producing brief representing, increasing the dimension for producing separable representation

Overall, in this extracting step, the appropriate attribute for signature recognition is extracted and pattern classes are formed if necessary.

C. Learning

In the next step that is neural network learning, as mentioned earlier, a multilayer Perceptron network with a post diffusion learning rule is used that is a comprehensive method for classifying data. It can be said that learning is a kind of learning by a supervisor. In this method it's possible that there is a relationship between the outputs and the weights or the errors are diffused from the output layer to the input one and the weights are corrected.

The main purpose is to plan a network that is initially educated by the existed learning data. Then by using the input vector to the network it can recognize its class. Such a network is widely used for pattern recognition tasks. In the next step, the difference between the signatures and the reference signatures are evaluated. If the difference is more than the predefined threshold, the signature is rejected.

IV. RESULTS

The experiment in this paper has been performed in a computer with following characteristics: Processor: Intel Core™ i5 CPU M480, Installed Memory (Ram):2.00 GB, Win 7 Home Premium, Software: Matlab R2014b.

In the proposed experiment, the database of signatures related to the first competitions of dynamic signature conforming (SVC 2004) is used. This database involves a set called Task1. The information of the signature attributes are shown in it as well as the time of its record and the pen position (pen-up and pen-down). The signatures in this database are related to 40 people and there are 40 samples for each person. From these 40 samples, 20 first samples are the major signatures and the other 20 samples are the secondary signatures signed by the professional forgers. The file is in a text format and is named $U_xSy.txt$. x is the number of user and it is ranged from 1 to 40. Signatures 1 to 20 are real and 21 to 40 are forged.

Data of these text files started with some samples and have 4 columns of: x , y , time (t) and the pen position (pen up=0 and

pen down=1). A figure of text file is shown in the following figure.

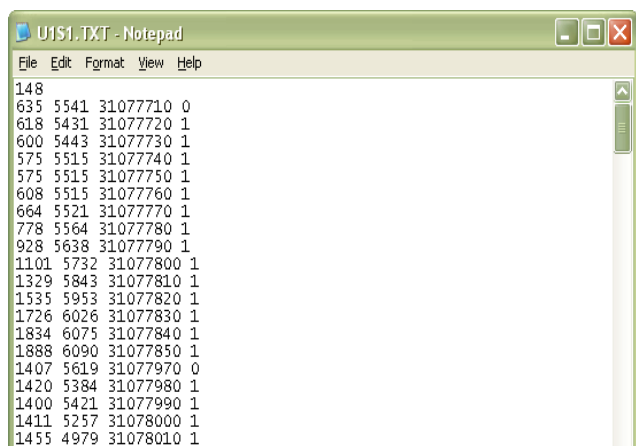


Figure 4. a section of text files U1S1.txt

For assessing the efficiency of the signature recognition methods, it is common to use a set of real signatures and forged signatures. By the use of signature database, false acceptance rate (FAR) and false rejection rate (FRR) are calculated.

False acceptance rate means that we accept someone forger as a person who has signed it and false rejection rate means that we reject someone who has signed it. For an ideal system, both false acceptance rate and false rejection rate should be a small number. False acceptance rate and false rejection rate depends on each other and decreasing one of them, increases the other. In practice, Equal error rate is used in evaluating system. Equal error rate is a point that false acceptance rate and false rejection rate equal each other. The chart of ERR calculation is shown here:

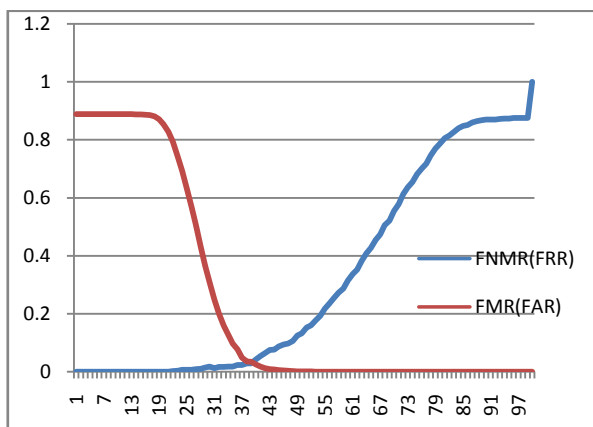


Figure 5. the chart of ERR calculation

In above chart, the range 0-100 shows the threshold amounts in horizontal axis and FAR and ARR in vertical axis. As it is clear, both positive false acceptance rate and negative false rejection rate are shown for all the samples and in two charts. The cross of these two charts shows equal error rate (ERR). This is a point in with the best amount of threshold is located on it. It should be explained that depending on the

TABLE I
COMPARISON OF THE PROPOSED METHOD WITH OTHERS METHODS

Method	FAR	FRR	Accuracy
Proposed method	13.92	15.69	86.74
Method of Livcki etal[7]	20.88	21.47	78.89
GA-SVM method[8]	15.63	17.78	83.60

identity recognition application and the safety, another point can be regarded as the threshold. The results of comparison are shown in the table I.

V. CONCLUSION AND FUTURE WORKS

In this paper, a method is proposed to recognize the handwritten signatures as the real or forged ones. Because of the increase in using the signature in financial documents and forging the signatures, Systems are developing to recognize the signatures by high accuracy and in the lowest time. This kind of systems should do a set of actions on the handwritten signatures. For example, they should do preprocess, extract attributions and form a system for accepting the real or forged signatures. 70 percent of experiment was for learning and 30 percent was for test. The results of proposed method showed that this model is able to recognize signature by a high accuracy and speed. It is proposed to study and research following items:

- Mail systems in companies and offices to send them without the possibility of being forged.
- Using special cameras in financial institutes for paying Cheque with determined amounts by controlling the date, price and signature.
- Signature recognition by biometric systems in offices and formal organizations such as document recording and ID recording companies.

REFERENCES

- [1] Maiorana, E., Campisi, P., Fierrez, J., Ortega-Garcia, J., & Neri, A. "Cancelable templates for sequence-based biometrics with application to on-line signature recognition" IEEE Transactions Systems, Man and Cybernetics, Part A: Systems and Humans, 2010, 40(3), pages 525-538.
- [2] Kolhe, S. R., & Patil, P. M. "Dynamic time warping based static hand printed signature verification. Journal of Pattern Recognition Research, 1, 52-65, 2009.
- [3] SAEED K., ADAMSKI M., "Extraction of Global Features for Offline Signature Recognition", Image Analysis, Computer Graphics, Security Systems and Artificial Intelligence Applications, WSFiZ Press, pp. 429-436, 2005.
- [4] Gahi, Y., Lamrani, M., Zoglat, A., Guennoun, M., Kapralos, B., & El-Khatib, K. "Biometric identification system based on electrocardiogram data". In New Technologies, Mobility and Security, NTMS'08., November 2008, pages 1-5.
- [5] Jung Wang, Xiao F. Liao and Zhang Yu, "Signature Recognition and Verification with Artificial Neural Network Using Moment Invariant Method", Second International Symposium on Neural Networks, Chongqing, China, May 30 - June 1, 2005. doi:10.1007/11427445_31
- [6] Sonkamble, S., Thool, D. R., & Sonkamble, B. "SURVEY OF BIOMETRIC RECOGNITION SYSTEMS AND THEIR APPLICATIONS", Journal of Theoretical & Applied Information Technology, (40)3, pages 45-51, 2010.
- [7] Liwicki, M., Malik, M. I., Alewijnse, L., van den Heuvel, E., & Found, B. "ICFHR 2012 Competition on Automatic Forensic Signature Verification (4NsigComp 2012)". International Conference Frontiers in Handwriting Recognition ICFHR2012, pages. 823-828. doi: 10.1109/ICFHR.2012.217

- [8] Kour, J., Hanmandlu, M., & Ansari, A. Q. "Online signature verification using GA-SVM". International Conference on Image Information Processing (ICIIP), November 2011, pages. 1-4.
- [9] Debnath Bhattacharyya, and Tai'hoon Kim, "Design of Artificial Neural Network for Handwritten Signature Recognition", International journal of computers and communications, Issue 3, Volume 4, pages 59-66, 2010.
- [10] Ali Karouni, Bassam Daya, Samia Bahlak, "Offline signature recognition using neural networks approach", Procedia Computer Science Volume 3, Pages 155-161, 2011. doi:10.1016/j.procs.2010.12.027
- [11] S.Odeh, M.Khalil, "Offline Signature Verification and Recognition: Neural Network Approach," IEEE, 2011, 978- 1-61284-922-5/11.



Ali Habiboghli received the B.S. degree from department of engineering of Islamic Azad University, KHOY Branch in 2004. He received the M.S degrees from electronic, computer engineering and information technology from Islamic Azad University of Qazvin Branch, Iran in 2007. From 2007 to already is with the Islamic Azad University of KHOY Branch. His research focuses on

artificial intelligence, algorithms, biometric and image processing.



Ramin Dehghani received the B.S. degree from department of engineering of Islamic Azad University Khoy Branch in 2005. He received the M.Sc. degree in computer science from Islamic Azad University, Khoy Branch, Iran in 2016. From 2008 to already is with Tejarat Bank in West Azerbaijan, Iran. His research focuses on Neural Network, Biometric.

A Report on Using GIS in Establishing Electronic Government in Iraq

Ahmed M.JAMEL
Department of Computer Engineering
Erciyes University
Kayseri, Turkey

Tolga PUSATLI
Department of Information Technology
Cankaya University
Ankara, Turkey

Abstract—Electronic government initiatives and public participation in them are among the indicators of today's development criteria for countries. After the consequent of two wars, Iraq's current position in, for example, the UN's e-government ranking is quite low and did not improve in recent years. In the preparation of this work, we are motivated by the fact that handling geographic data of the public facilities and resources are needed in most of the e-government projects. Geographical information systems (GIS) provide the most common tools, not only to manage spatial data, but also to integrate with non-spatial attributes of the features. This paper proposes that establishing a working GIS in the health sector of Iraq would improve e-government applications. As a case study, investigating hospital locations in Erbil has been chosen. It is concluded that not much is needed to start building base works for GIS supported e-government initiatives.

Keywords—Electronic government, Iraq, Erbil, GIS, Health Sector.

I. INTRODUCTION

Electronic government or e-government in short, is defined in many research articles and reference sources; more-or-less the definition is similar to what is given in [1], where the author defines it as application of Information and Communication Technology (ICT) in the public sector to improve the efficiency, effectiveness, transparency and accountability of government.

In this paper, we represent our findings of our research on e-government transition case study on health services in Erbil, Iraq [2]. Basically, the research concentrates on health centers in Erbil to demonstrate benefits of employing GIS in e-government initiatives and to discover potential and possible obstacles in establishing and using such an information system that can/will be used as a part of infrastructure of e-government. The research question addressed here is: Can Iraq's local governments establish a working GIS to handle hospital distributions and resources to speed up e-government initiatives in its health services?

Before reporting the outcome of the research, it is beneficial to give a brief background on the subject.

II. BACKGROUND

Basically, e-government is based on citizens accessing information and services provided by the government 24 hours a day and 7 days in a week through the Internet [3]. One of the aims of e-government initiatives is to provide an easier way to enhance the mechanism of government. Such initiatives include centralization and management of text based and spatial data, such as health centres.

E-government applications provide easy, fast and digital interaction between citizens, businesses and government itself among its departments. E-government represents a "transition" from routine management style to a contemporary way. Related literature prefers "transition" instead of "change" since they are different in the way they occur. Transition is meant to include improving an old situation, to a new one in a controlled manner while, "change" occurs naturally without any control.

Basically, there are four types of e-government delivery models classified according to the services that the government provides and based on the nature of the parties involved: Government to Government (G2G), Government to Business (G2B), Government to Citizens (G2C) and Citizen to Citizen (C2C).

As a directive and authority, European Commission prescribes 8 services that government provides to businesses [4]; such services are presented to public and private sectors regardless their size. These services are:

- Social contributions for employees.
- Corporate tax.
- VAT.
- Registration of a new company.
- Submission of data to statistical offices.
- Customs declarations.
- Environment-related permits.
- Public procurement.

Similarly, same commission prescribes 12 services that government has to provide to citizens [4]:

- Income taxes.
- Job search services.
- Social security benefits.
- Personal documents (passports / driver's license).
- Car registration.
- Application for building permission.
- Declaration to the police.
- Public libraries.
- Certificates.
- Enrolment in higher education/university.
- Announcement of moving.
- Health related services.

For Example: instead of going to the tax office and standing in line, the citizen is now able to pay a tax through the government's tax website, whenever he wants, even at night from his home. Another service is in the health sector. When a patient goes to doctor for the first time, he can register his name through a centralized system. If his doctor wants a blood test, MRI or X-Ray, then directly his name will appear in the system in that department. Furthermore, he does not need to carry his results with him. The results will appear in his doctor system, directly.

Iraq is trying to catch up with the rest of the world in most of the industries after surviving two wars in 15 years. Within this nationwide development, it is not surprising to see that there will be a controlled passage in managing governmental offices from paper based to electronic media via electronic government initiatives.

However, such a transition is too big to handle by a single government body even by a ministry. Similar attempts have been monitored in many countries e.g. Europe (<http://ec.europa.eu/dgs/informatics/ecom>), USA (<http://www.usa.gov>), Turkey (<http://www.turkiye.gov.tr>) and South Korea (<http://korea.go.kr>). A common observation shows that those attempts are done through a series of actions by multiple parties. These parties include local governments, ministries and central government, as well as citizens.

After the physical infrastructure of Iraq was destroyed, the technical infrastructure began to redevelop. Some countries experienced worse than the Iraq situation yet could stand and build their infrastructure and in the process becoming one of the best countries in terms of technology and e-government services. A good example is South Korea, while lesser case examples include Angola and Uganda, where civil conflicts occurred. Meanwhile their states are working to transform paper-based processes to digital processes through e-government initiatives [5].

Transforming into digital government requires a lot of steps and may take years.

Another example is Turkey, where serious steps have been taken in the country where both central and local governments have established e-government services for both citizens and businesses in the last decade.

From these examples about countries, we conclude that building e-government in developing countries is possible, step-by-step. By taking the experience of these countries into consideration and learning the obstacles they faced during establishing e-government, we can see that e-government is present in developing countries, as well.

In this study, Erbil city has been taken as an example of establishing e-government in the health sector. The health sector is an important sector and government has a huge role in providing services to citizen in this sector. The City Erbil (Hawler) is located in the north of Iraq in Erbil state. It has a surface area of 13,165 km². Apart of the city center, the districts of Dashti Hawler, Makhmour, Koya, Shaqlawa, Soran, Rawanduz, Choman and Mergasor are populated areas. The estimated population of Erbil is 1,612,692 as of 2011 [6]. It is the fourth largest city in Iraq after Baghdad, Basra and Mosul.

III. EXISTING E-GOVERNMENT PROGRAMS IN IRAQ

The first serious plans for establishing e-government in Iraq were in June 2014 when United Nations called on member states to help the new Iraqi government for institution building. The project was initiated by Italian Minister for Innovation and Technologies and the Iraqi Minister of Science and Technology. In order to initiate an e-government project, the Italian government provided technical and financial assistance to the Iraqi government. The plan linked Ministries via an Intranet. United States Agency for International Development (USAID) in cooperation with Iraq Ministry of Science and Technology also put a strategy in place from 2007-2010 to develop Iraqi electronic government project [7].

Currently, e-government in Iraq is developing step by step. The Iraqi portal is <http://www.egov.gov.iq/egov-iraq/index.jsp>. This portal provides simple services to citizens, businesses and government. For example; downloading passport application forms <http://www.iraqinationality.gov.iq/eForm>, booking flight tickets <http://ia.com.iq/>, downloading driving license application forms <http://www.itp.gov.iq/lic1.php>, Citizens affairs <http://citizenaffairs-egov.com/>, monitoring traffic fines http://itp.gov.iq/fines_ar.php and online application to universities in Kurdistan region <http://www.regayzanko.com/>.

IV. E-GOVERNMENT OBJECTIVES AND REQUIREMENTS

Shifting paper based work on to electronic media has several considerable benefits, not only for government, but also the public and private sectors.

Firstly, it is fast: instead of going to the government departments, a person can fulfill their work in a faster way electronically. For example: having an appointment from a hospital over the Internet. The patient interacts with the hospital from a single point and looks for eligible doctors to get diagnosed. As the person does not need to be present at the transaction place physically, time required to stand in line is saved as they can fulfill their work by interaction with government via website. For example: in a document service such as a passport, a person can submit their scanned personal documents via the Internet, in this case he or she will save a time by not going to passport office.

Doing the paperwork electronically reduces additional costs for the transaction as well. Instead of going to the government department and paying for transport, a person can fulfill her work while at home. For example: a person can enroll through the Internet to a university avoiding transportation cost.

Although minimum requirements are necessary to get involved in e-government initiatives, the benefits of such initiatives increase computer literacy as well. People understand and learn how to use the Internet, websites and learn how to fulfill their work digitally. In this case, the efficiency and awareness of people will increase. The technology will force people to learn how to use Internet and adapt to modern life. The e-government process and information will be more accurate than traditional paper base models, because it is easy to control. For example: the electronic fulfilling is more accurate than that made manually [5].

There are considerable examples in practice to justify e-government objectives. Such a list includes but not limited to the following items as discussed in [5] and [8].

E-government provides quick services to local citizens with less cost, and enables them to practice democracy, thus reducing the time required to complete transactions. Achieving effective communication and reducing the complexity of administration improves the level of services to citizens and institutions. It can also overcome the errors that employees may make in a manual system.

Creating a better working environment using ICT in the institutions and the establishment of an infrastructure for e-government, helps to improve the communication interface between the government and employees. Examples such as the tracking and tracing of service provision, the ability to conduct services in steps and the indication of time duration for service completion, all promote accountability, transparency and satisfaction, while decreasing administrative load.

There are many requirements for applying e-government initiatives, including technological, management, legal and human considerations. To build an effective e-government system, a state must develop a strategy by a qualified team to ensure technical, organizational and legal infrastructure. For applying e-government we have compiled following fundamentals and requirements from [5] and [9].

Network: for e-government employees to do their duties and interact with each other, their computers have to be connected on a network to share data and information. Additionally they have to learn about using computer, Internet and e-government application. Furthermore they have to have certificate provided by government that allow them to work on e-government application.

Improve government organizations and institutes: to build an effective e-government we have to improve government institutes and organizations through changing the regulatory mechanism of government. Traditional mechanisms are not suitable with modern, fast and flexible e-government.

Improvement of employee for dealing with e-government: in order to train government employees to be part of an e-government environment, the government must prepare training courses.

Awareness of citizens: the citizens have to know about using a computer and the Internet while having to be acclimating with the digital environment. The citizens can learn through basic computer courses and also online courses with tutorials via the Internet.

V. GIS

GIS are computer systems that collect, edit, input, retrieve, store, analyze and output spatial and non-spatial data for specific purposes [10]. GIS is able to input geographic information (maps, aerial images, etc), data (names and tables), process (process and review errors), store, analyze (spatial and statistical) and view on a computer as reports, maps or as charts.

Such systems are useful during engineering planning, design, management of distribution for hospitals [11].

The GIS components are a set of parts including network, computer hardware, computer software, data and information. These components allow GIS to perform interrelated tasks [10]. The network is the infrastructure to GIS that enhance the reusability and accessibility of geo-referenced data and analysis tools [11].

In order to establish and build an e-government project in the health sector using GIS in a developing country, three steps are recommended.

1. collecting clean data by which we mean that the data about hospitals should be accurate and clear. For example the number of hospitals, coordinates of each hospital, the number of doctors, nurses, MRI and Tomography systems and beds etc.
2. recording the data in a central database and sharing it through network servers. For example, creating a database and saving the collected data, then putting a database on a server and sharing it with responsible departments in government.
3. updating the data timely. For example; the number of doctors and beds may change or a new hospital may be built or one may be closed. These changes in the data should be updated promptly in the database for accuracy.

After fulfilling these steps, we are able to determine for example, the shortest path for ambulances. In this way, a successful GIS project can be established.

VI. Outcome

Iraq's position in the development ranking can be promoted by e-government. This can be supported by managing spatial data of the public facilities and resources. Strong discussions should be started and/or accelerated on this subject in local and central governments in Iraq.

Budgets should be prepared for buying hardware and software, while the data collection should be started with clean and timely data. Obviously, data collection and reducing variety in the data has priority.

Future works can and should be established to keep the topic relevant and on the table. Promoting e-government initiatives will enhance public services in Iraq. This enhancement will be good in saving valuable time that is wasted in data duplication in the government side and will avoid unnecessary waiting time at the service points.

Strong project managements should be carried out to collect and maintain spatial and non-spatial data on the public facilities as well as human resources.

As a complementary study, the digital gap can be studied so that authorities can focus on specific geographic areas and people; in this way, the usage of the e-government application by the public, hence participation can be ensured.

VII. CONCLUSION

E-government initiatives require relatively new technology-and-management supported transformation strategies that governments use to connect their organizations to each other and provide services to citizens and businesses. During the preparation of this research, we have learnt that managing spatial and non-spatial data is required in most of such services.

Recalling the research question, "can Iraq's local governments establish a working GIS to handle hospital distributions and resources to speed up e-government initiatives in its health services", this research yields positive and encouraging results. Only by collecting limited spatial and non-spatial data, the researcher was able to build a basic GIS to answer simple questions such as "where is the closest hospital within a 2km radius" or creating thematic maps for enhanced reporting. We have seen that setting up a GIS in Erbil does not require extreme technical knowledge; however, a strong limitation that obtaining clean and up-to-date data is an important concern. Although it is possible to establish a GIS in this sector, the starting point should not be acquiring hardware and software, but collecting clean data from the field and ensuring that the data gets updated timely. Additionally, examples from the literature show that conflicts in the countries are causing big problems, however this may not prevent governments investigating possibilities for establishing e-government initiatives.

Thus, Iraq can certainly start designing projects for such initiatives and the health sector can be one of the areas to begin with.

As underlined in the literature, common criteria and basics to assess performance of e-government programs and the services are data integrity and purity, mainly.

REFERENCES

- [1] M. Yildiz, "E-devlet ders notları", 2011
- [2] A. M.Jamel, "An investigation of electronic government transition case study in health: hospitals in Iraq", Master Thesis, Cankaya University, 2013
- [3] Organization for Economic Co-Operation and Development OECD), "E-government for better government. OECD e-government studies", 2005, p.206
- [4] B. Lörincz et al "E-government benchmark framework 2012-2015", 2012
- [5] D. Jameel, "E-government and the barriers for applying it". Islamic University of Economic and Administrative Studies, 2012, p.38
- [6] Central organization for statistics, "Population projections by governorates & social origin 2010", 2011, available at <http://www.cosit.gov.iq/AAS13/population/pop%2812%29.htm>
- [7] M. Al-dabbagh, "Electronic government in Iraq: challenges of development and implementation", Swedish Business School at Örebro University, 2011
- [8] B. Lörincz et al, "Digitizing public services in Europe: putting ambition into action 9th benchmark measurement" 2010
- [9] M. Yildiz, "E-government research: reviewing the literature, limitations, and ways forward". Government Information Quarterly. 24, 2007, p.19
- [10] M. Demers, "Fundamentals of geographic information systems" 2009
- [11] S. Aziz, "The benefits and potential implementation problems of GIS in the water distribution services of municipalities case study: the Kirkuk water directorate", Master Thesis, Cankaya University, 2012

Satellite Image Classification by Using Distance Metric

Dr. Salem Saleh Ahmed Alamri
Department of Engineering Geology ,
Oil & Minerals Faculty, Aden University,
Aden, Yemen

Dr. Ali Salem Ali Bin-Sama
Department of Engineering Geology ,
Oil & Minerals Faculty, Aden University,
Aden, Yemen

Dr. Abdulaziz Saleh Yeslam Bin –Habtoor
Department of Electronic and Communication Enigeerinng,
Faculty of Engineeninng & Petrolem,Hadramote University,
Mokula , Yemen

Abstract— This paper attempts to undertake the study satellite image classification by using six distance metric as Bray Curtis Distance Method, Canberra Distance Method, Euclidean Distance Method, Manhattan Distance Method, Square Chi Distance Method, Squared Chord Distance Method and they are compared with one another, So as to choose the best method for satellite image classification.

Keyword: *Satellite Image, Classification, Texture Image, Distance Metric,*

I. INTRODUCTION

Defines image classification as particular class of pattern recognition. Classifiers are described under board categories such as supervised and unsupervised classifiers, parametric and non-parametric, fuzzy classifier and knowledge base classifiers [1].

Defines three major steps involved in the typical supervised classification procedures as follows:

- **Training Stage:** The analyst identifies reprehensive training area and develops a numerical description of the spectral attributes of each land cover type of interest in the scene.
- **Classification Stage:** Each pixel in the image is categorized into the cover class it most resembles. if the pixel is not matching to any predefined class then it is labeled as unknown.
- **Accuracy Assessment:** The classified image is compared with some reference image or ground truth to check the accuracy of the classification [2].

They are proposed an algorithm for very high-resolution satellite image Classification that combines non-supervised segmentation with a supervised Classification and the result

show very good performance of approach in comparison to existing techniques [3].Minimum distance classification method in satellite Image is a simple and quick method that does not include covariance They are proposed an algorithm for very high-resolution satellite image Classification that combines non-supervised segmentation with a supervised Classification and the result show very good performance of approach in comparison to existing techniques [3].Minimum distance classification method in satellite Image is a simple and quick method that does not include covariance information and Maximum likelihood classification method is widely used in remote sensing image and can be regard as one of the most reliable techniques [4].They are proposed a new algorithm for texture classification based on logical operators is presented. Operators constructed from logical building blocks are convolved with texture images. This algorithm is applicable to different types of classification problems which is demonstrated by segmentation of remote sensing images, compressed and reconstructed images and industrial images[5]. He proposed a simple scheme which used local linear transformations and energy computation to extract texture features. This simple scheme often gives good results but is not consistent in performance. The statistical methods share one common weakness, of primarily focusing on the coupling between image pixels on a single scale and are also computationally intensive processes [6].He proposed texture segmentation and classification for texture features images based on the grey level co-occurrence probabilities (GLCP) [7].They is proposed classification method based on the Discrete Cosine Transform (DCT) coefficients of texture images by used two popular soft computing techniques namely neuron computing and neuron-fuzzy computing [8].He is proposed to perform unsupervised image classification based on texture features by using a novel evolutionary clustering

method, named manifold evolutionary clustering MEC according to a novel manifold-distance-based dissimilarity measure, which measures the geodesic distance along the manifold [9]. They are present a new Euclidean distance for images, which we call Image Euclidean Distance (IMED). Unlike the traditional Euclidean distance, IMED takes into account the spatial relationships of pixels. The key advantage of this distance measure is that it can be embedded in most image classification techniques but it has one limitation of IMED is that it does not always provide the best recognition result comparing with some other intelligent metrics under the nearest neighbor rule. [10]. The conventional classification techniques typically measure closeness in Euclidean space. This leads to an anomaly that two similar classes of vegetation having an identical vegetation index profile with different growing practices (sowing, senescence, harvest,.... etc.) appear as points separated by a large distance leading to different class labels [11].

II. SATELLITE IMAGE CLASSIFICATION

Image classification refers to the task of extracting information classes from a multiband raster image. The resulting raster from image classification can be used to create thematic maps. Depending on the interaction between the analyst and the computer during classification, there are two types of classification: supervised and unsupervised

a. Supervised classification

Supervised classification uses the spectral signatures obtained from training samples to classify an image. With the assistance of the Image Classification toolbar, you can easily create training samples to represent the classes you want to extract. You can also easily create a signature file from the training samples, which is then used by the Spatial Analyst multivariate classification tools to classify the image.

b. Unsupervised Classification

Unsupervised classification finds spectral classes (or clusters) in a multiband image without the analyst's intervention. The Image Classification toolbar aids in unsupervised classification by providing access to the tools to create the clusters, capability to analyse the quality of the clusters, and access to classification tools.

III. TEXTURE CLASSIFICATION

Texture classification is an image processing technique by which different regions of an image are identified based on texture properties. This process plays an important role in remote sensing applications which is depending on texture feature extraction of the images. This paper using the texture classification to classifiers of satellite image with distance metric methods.

VI. DISTANCE METRIC

Distance metric is the method using in image classifications for examining the performance of textural features of image. Distance metric are two types using in the image classification minimum distance and maximum distance in this paper will used a minimum distance classifier based on the assumptions of patterns and classes of the images by the Euclidean distance [5].

IV. DISTANCE METRIC METHODS

This paper discusses six distance metric methods

:

A. Bray Curtis Distance Method

Bray Curtis distance methods can be expressed by this equation:

$$d_{BC}(x, y) = \sum_{i=1}^d (|x_i - y_i| / (x_i + y_i)) \quad (1)$$

Where:

$d_{BC}(x, y)$ is the Bray Curtis Distance in the 2-dimensional x, y metric distance.

B. Canberra Distance Method

Canberra distance methods can be expressed by this equation:

$$d_C(x, y) = \sum_{i=1}^d (|x_i - y_i| / (|x_i| + |y_i|)) \quad (2)$$

C. Euclidean Distance Method

Euclidean distance methods can be expressed by this equation:

$$d_E(x, y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2} \quad (3)$$

D. Manhattan Distance Method

Manhattan distance methods can be expressed by this equation:

$$d_M(x, y) = \sum_{i=1}^d |x_i - y_i| \quad (4)$$

E. Square Chi Distance Method

Square Chi distance methods can be expressed by this equation:

$$d_{Chi}(x, y) = \sum_{i=1}^d (x_i - y_i)^2 / (x_i + y_i) \quad (5)$$

F. Squared Chord Distance Method

Squared Chord distance methods can be expressed by this equation:

$$d_{sc}(x, y) = \sum_{i=1}^d (\sqrt{x_i} - \sqrt{y_i})^2 \quad (6)$$

VI. EXPERIMENTS VERIFICATION

A. Testing Proceeding

The classification was implemented using (MATLAB R2007a, 7.4a) and classified satellite image by using Six Methods of Distances Metric, Bray Curtis Distance Method, Canberra Distance Method, Euclidean Distance Method, Manhattan Distance Method, Square Chi Distance Method, and Squared Chord Distance Method in 10 texture images illustrated on the Fig(1).

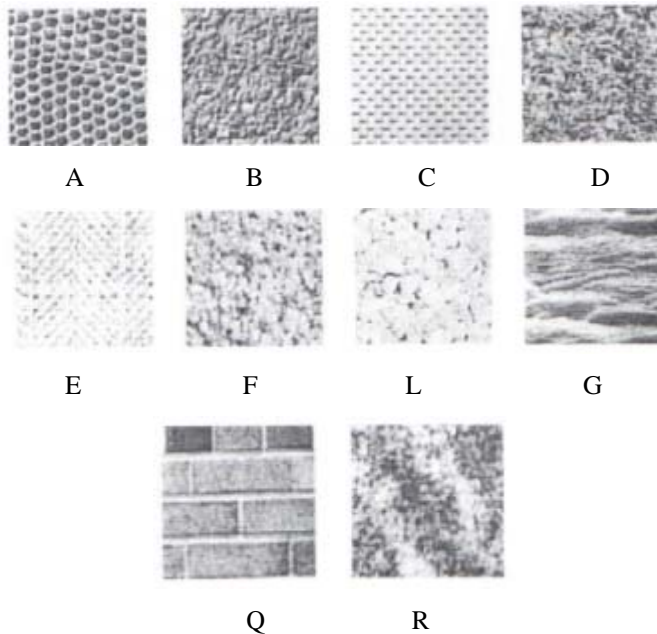


Fig.(1)Textures Image Used

B. Simulation Results

The simulation Results are shown in the Fig.(2).



Fig.(2)Simulation Result Image

VII. CONCLUSION

In this paper applied six distance methods as Bray Curtis Distance Method, Canberra Distance Method, Euclidean Distance Method, Manhattan Distance Method, Square Chi Distance Method, and Squared Chord Distance Method to classified satellite images based on the texture classification Methods to texture feature extraction of the images applied in the ten texture image. The important conclusions that can be drawn from the present analysis as follows:

- 1-Performance of Bray Curtis and Canberra methods are better than other methods, classification accuracy is 85%.
- 2-Performance of Square Chi and Squared Chord are better than Euclidean and Manhattan distance, classification accuracy is 76%.
- 3-Performance of Euclidean and Manhattan distance are less than other methods, classification accuracy is 71%.

REFERENCES

- [1] R. C. Gonzalez, R. E. Woods, "Digital Image Processing" Pearson Education, 2004.
- [2] A. Bharti, "A Decision Tree Approach to Extract Knowledge for Improving Satellite Image Classification", international institute for Geo-information Science and Earth Observation, Enschede, The Netherlands, and Indian institute of remote Sensing, National Remote Sensing Agency (NRSA), Department Space, Dehradun, India, PP.7, 2004.
- [3] X. Gigandet, M. Bach Cuadra, A. Pointet, L. Cammoun, R. Caloz and J.Ph. Thiran, "Region Based` Satellite Image Classification: Method and Validation", IEEE 2005.
- [4] M. Hosseini, M. R. Saradjian, A. Javahery and S. Nadi, "Noise Removal from Land Cover Maps by Post processing of Classification Results", IEEE 2007.
- [5] Vidya Manian, Ramon Vasquez, Praveen Katiyar, "Texture Classification Using Logical Operators", IEEE Transaction on Image Processing, Vol. 9, No. 10, pp.1693-1703, 2000.
- [6] W. K. Pratt, "Digital Image Processing", New York: Wiley, 1991.
- [7] Hong-Choon Ong, Hee-Kooi Khoo, "Improved Image Texture Classification Using Grey Level Co-occurrence Probabilities with Support Vector Machines Post-Processing", European Journal of Scientific Research .Vol.36 No.1, pp.56-64, 2009
- [8] Golam Sorwar, Ajith Abraham, "DCT Based Texture Classification Using A Soft Computing Approach, Malaysian Journal of Computer Science, Vol. 17 No. 1, pp. 13-23, 2004
- [9] Maoguo Gong, Liefeng Bo, Ling Wang, "Image texture classification using a manifold distance-based evolutionary clustering method", Optical Engineering 47_7_, 077201 _July 2008
- [10] Liwei Wang, Yan Zhang, Jufu Feng, "the Euclidean Distance of Images", NNSF (60175004) and NKBRSF, (2004CB318005), Center for Information Sciences School of Electronics Engineering and Computer Sciences, Peking University Beijing, 100871, China
- [11] Sudhir Gupta , K. S. Rajan , "TEMPORAL SIGNATURE MATCHING FOR LAND COVER CLASSIFICATION", International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Volume XXXVIII, Part 8, pp.492-497, Kyoto Japan 2010.

AUTORS PROFILE



#1 Dr. Salem Saleh Ahmed Alamri: Received the B.E degree in Mechanical Engineering, University of Aden, Yemen, 1991, M.Sc. degree Computer science (IT) from North Maharashtra University (N.M.U), India, Jalgaon, 2006, PhD degree in computer science (S.R.T.M.U), India, Nanded., 2011.

He has 12 international papers, 8 international conferences. assist professor HOD of department of engineering geology in Minerals & Oil Faculty, Aden University, Yemen; He is membership in International Association of Engineers (IAENG)



***2 Dr. Ali Salem Ali Bin Sama:** Received the B.E degree in Computer Engineering, Balqa University, Faculty of Applied Engineering, Jordan, 1997, M.Sc. degree in Computer science from University of Science, Malaysia, 2006, PhD degree in computer Science,

Malaysia, 2011, He has 11 international papers, 5 international conferences. He is assist. professor of department of engineering geology in Minerals & Oil Faculty, Aden University, Yemen, Assistant Professor in a Department of Computer Science and Information, College of Sharia and Islamic Studies at Al-Ahsa, Imam Muhammad bin Saud Islamic University, in Saudi Arabia, From 25/08/2014 until now.



****3 Dr. Abdul-Aziz Saleh Ysalem Bin-Habtoor:** Received the B.E degree in Electrical and Electronic Engineering, University of Aden, Yemen, 1991, M.Sc. degree Electronic and Communication from University of Technology, Baghdad, 2000. PhD degree in Electronics and

Communication Engineering Narrow Specialization in Computer Networks and Data Communication. From University of Technology, Baghdad, 2004. He gain work experiences as follows: the current job is asses. prof. at Faculty of Sciences and Arts at Sharourah, Najran University, KSA since 2010 to Now 2016. General Administrator of Information Networks in Hadhramout University of Science and Technology from 2008 to 2010. Coordinator of Hadhramout University with ministry of Higher Education especially in the information networks from 2008 to 2010. Assist Professor Electronic and Communication Engineering, Faculty of Engineering and Petroleum, Hadhramout University, Yemen.

Cybercrime and its Impact on E-government Services and the Private Sector in The Middle East

Sulaiman Al Amro

Computer Science (CS) Department
Qassim University,
Buraydah, Qassim, 51452, KSA

Abstract—This paper will discuss the issue of cybercrime and its impact on both e-government services and the private sector in the Middle East. The population of the Middle East has now become increasingly connected, with ever greater use of technology. However, the issue of piracy has continued to escalate, without any signs of abating. Acts of piracy have been established as the most rapidly growing (and efficient) sector within the Middle East, taking advantage of attacks on the infrastructure of information technology. The production of malicious software and new methods of breaching security has enabled both amateur and professional hackers and spammers, etc., to target the Internet in new and innovative ways, which are, in many respects, similar to legitimate businesses in the region.

Keywords—cybercrimes; government sector; private sectors; Middle East; computer security

I. INTRODUCTION

The evolution of Information Technology has given rise to the wide use of cyberspace, in which the Internet provides equal opportunities for access to information. Due to this increased use, the misuse of technology has also continued to increase, leading to an increase in cybercrime. Cybercrime refers to an unlawful act, in which the computer is either the tool or the target, or both. Cyber Security refers to the mechanism by which computer-based equipment, information and services are protected from illegal and unauthorized access. By March 2016, the Middle East had achieved 3.3% of the total worldwide use of the Internet [11, 22], having increased by 1648.20% between 2000 and 2009, in contrast to a global increase of 259.6% over the same period. This large number of users in the Middle East has ensured that the Internet has become a popular means of communication, as well as opening up new opportunities for online business.

The Middle East has witnessed a rapid increase in Internet development. This current research will: Firstly, review the current level of protection of information resources of the e-government community in the Middle East. Secondly, there will be an investigation of the ways in which e-government is misused by attackers, including identifying the motivation behind cyber-attacks. Thirdly, there will be a discussion of the ways in which cybercrime impacts on e-government in the

Middle East. Finally, there will be a discussion of cybercrime as an emergent form of crime in the Middle East, and an examination of the ways in which the governments of the Middle East react to these offenses, thus leading to the lessons to be drawn for Internet use within the context of Middle East's economic, political, and legal conditions.

II. BACKGROUND

Cybercrime is a global issue, and no country can be isolated from its effects. With the continuous increase of Internet usage, there is a global increase in the rate of cybercrime. The Middle East has a high rate of Internet usage, which increasingly ensures that its online resources are at risk. A number of studies have attempted to resolve this issue by deploying new technology, or enforcing a number of policies, both inside and without, an organization. However, the issue continues to persist and increase. This study will investigate a number of areas related to this issue that have not previously received sufficient attention in the Middle East.

The literature includes a number of valuable studies discussing the issues and challenges of cybercrime. The following section will discuss the studies most relevant to the current research.

III. TAXONOMY

1. E-government

The term 'e-government' refers to electronic government, also known as Internet government, digital government, and online government. Connected government consists of the digital interactions between the following: (1) citizens and their government; (2) between governments and government agencies; (3) between government and citizens; (4) between government and employees; and (5) between government and businesses/commerce [6].

2. Cyberspace

Cyberspace consists of the notional environment in which communication over computer networks occurs.

3. Cybercrime

Computer crime, also known as cybercrime, consists of crime involving a computer and a network.

4. Cyber-Attack

Cyber-attack consists of any type of aggressive maneuver employed by individuals, or entire organizations, targeting computer information systems, infrastructures, computer networks, and/or personal computer devices by various means or malicious acts. These generally originate from an anonymous source, which steals, alters, or destroys a specified target by hacking into a susceptible system.

5. Computer Security

Computer security, is also known as cybersecurity or IT security, and consists of the protection of information systems from theft or damage to the hardware, and/or software, and to any information they contain, as well as from disruption or misdirection of the services they provide.

IV. IMPORTANCE AND PROBLEM

Technology is currently evolving exponentially, leading to an accompanying evolvement of software vulnerabilities, ensuring security is always at risk. Attackers exploit software vulnerabilities in order to carry out cybercrime. The application of policies, or deploying techniques, to combat cybercrime requires careful investigation into both present and potential threats. This becomes significant for an e-government sector providing a wide range of services for the public.

This current research will focus on the type of cybercrime relevant to each government agency, in a case-by-case manner, developing a framework of a solution to defend against such threats and avoid potential future threats.

Cybercrimes are increasingly affecting government organizations globally, being attractive to criminals due to the low cost involved and ease of access, as well as the low risk in comparison to physical crimes. Cybercriminals exploit Internet technologies wherever they can, taking advantage of millions of users of e-government services to cause rapid, and considerable, damage. The significant loss of resources, and the fact that it is time consuming to recover after such crimes, has led to this being a high priority security issue.

Government organizations may suffer from a lack of knowledge concerning the types of cybercrimes they currently face, and may face in future. Therefore, the development of a framework to defend and avoid such threats is needed.

Cybercrime is an attack on the information of both individuals and groups, which can prove extremely harmful, as criminals are aware that all organizations and government agencies, as well as the general public, rely on computers to save their documents. The following are general classifications of the types of attack.

1. Attack on individuals

Criminals offer a range of false promotions, tricking individuals into providing personal information. A further

crime consists of human trafficking, including the production of child pornography. Social networking sites and chat groups can also lead to serious cybercrime.

2. Attack on property

This includes harmful programs disseminated through Internet sites, e-mail or personal chat, which enable criminals to steal information easily from a computer system, as well as obtaining unauthorized access to the Internet.

3. Attack on the private sector

This occurs when criminals penetrate systems and companies whose servers store confidential and sensitive data. Hackers are able to access information, including that on company funds, which can result in significant issues.

4. Attack the on public sector

In a process known as 'cyber terrorism', criminals access the databases of government sectors, in order to use their information. This then reduces the effectiveness of a government and thus the faith of citizens in the government.

This research will address such issues associated with the attack on the public sector in the Middle East, and will establish a number of solutions.

V. LITERATURE REVIEW

A number of studies have been undertaken to establish the factors limiting cybercrime at the state level. [5] state that the study of social behavior can limit the spread of this phenomenon, on the premise that an individual more likely to undertake piracy is one who tends not to share with others, or show interest in them and their innovations, and tends to be isolated, spending a considerable amount of time in discussions over the Internet.

Liang et al. (2010) have examined the detection of electronics in China, in order to obtain a solution to such crimes. They have established that, although the use of the Internet has led to economic development in China, it has also led to a rise in electronic crime. They conclude that government needs to regulate the Internet, and administrative rules and regulations need to be put in place to cover electronic crimes (e.g. Internet child pornography and gambling) [1].

Nain and Patra [4] undertook a study on electronic crime in India. They suggest that, in order to guarantee success, preventive measures need to be in place at an early stage, along with an adequate protection plan. With the global increase in the use of the Internet, there is pressure for governments and organizations to establish appropriate rules and strategies for dealing with cybercrime, thus ensuring that individuals can act online without fear of exploitation.

The study by [2] is similar to the objectives of the study carried out by Akaroal [5]; however, their most important findings consist of the importance of new techniques relying on IQ in reducing these crimes.

The study by [3] aims to establish a solution for dealing with electronic crimes in different countries. Their study

recommends that these crimes can be most effectively overcome by being broadly classified into three categories: (1) electronic laws; (2) education; and (3) policy-making. These three laws can be given the facility to deal with cybercrime.

VI. DISCUSSION

Research is currently being undertaken into the issue of cybercrime, and its implications for The Middle East, particularly in the area of e-government services. Many recent studies have proposed efficient implementation techniques to enhance the detection of cybercrime. However, their inability to detect innovative, and unknown, attacks make them inappropriate for dealing with daily and new threats [12, 13].

A. *Examples of Cybercriminals' targets*

Cybercriminals currently attack popular sites in the region, including the large number of social networking sites. The majority of employees and individuals with access to the Internet currently use social networks. A number of studies have revealed that social networking can open a 'back door' to corporate information technology platforms, leaving companies and individuals at risk of losing information, as well as identity theft and other malicious attacks. Cybercriminals seek out sites with a large number of users who display poor security awareness, with social networking sites being an effective area. The Middle East currently has over twenty-seven local social networking sites, which can be used by hackers to infect users with malware or phishing sites, as well as to steal passwords and accounts and open security holes in the victims' machines.

Social networking sites are not secured in the Middle East, being sufficiently local to protect user privacy and sensitive information. Many new forms of Internet crime are performed on a daily basis while an individual is in the process of using social networks. Users in the Middle East also benefit from international social networks (e.g. Facebook and MySpace, friendships, blogging and other activities) that, if not approached with great caution, can lead to identity theft and malicious activities against home users, as well as (should there be no relevant policies in place) workers in the private and public sectors. The risk is very high, not only in social networks, but also in networks of peer to land, Web 2.0, chat and popular applications. This leads to the need for security awareness training in the Middle East.

A further important factor is that the majority of Middle Eastern countries currently face issues of unemployment. The numbers are increasing daily, and, if not taken into account, will affect the growth of cybercrime in the region. The World Bank states that: "the Middle East and North Africa countries are facing major challenges. They have to work itself to generate 100 million new jobs by 2020, or instability in the region will increase" [13]. Statistics reveal that the unemployment rate is particularly high amongst young people in the region, primarily University graduates with proficiency with computers and the Internet. In addition, Internet access is widely available throughout the region at the lowest rates online, by means of Internet cafes, leading to easy access, even for those without Internet access at home [14]. All these factors

combine to create a new generation of local hackers and cybercriminals. The majority focus on financial reasons, despite some having terrorist motives. Many are not aware of deep programming, i.e. hackers who are able to create their own malware or viruses, but take advantage of the many free sites in Arabic that can help them understand the basics behind hacking techniques, with links to underground hacking websites in foreign languages, and even obtaining free tools. These 'script kiddies' are now the greatest threat in the Middle East, due to a combination of having time on their hands, and the low cost of Internet access and Internet cafes, which can be employed to launch their attacks with ease.

Many international organizations warn that the Middle East has become a major source of cybercrime, i.e. Saudi Arabia, as the leading country in the region, is a target, as well as the source of malicious activity online, being Number 38 throughout the world. Saudi Arabia is also the primary source of malicious attacks on the Gulf Cooperation Council (GCC). Egypt has been identified as one of the most fraudulent countries globally, with approximately 1763 instances of fraud, closely followed by Saudi Arabia, the UAE and Qatar. Thus, it can be seen that electronic crimes are increasing in the region, due to the growth of the user base with weak security awareness and the absence of regulations. However, even traditional electronic crime (i.e. Phishing) has unique properties in the Middle East. Due to the religious and political issues in the region, hackers are able to successfully send political or religious messages, urging users to open e-mail attachments, and thus infecting the computer with malware to attack infrastructure targets in the Middle East, including e-commerce, banking, communications and government [10].

A further reason for the Middle East (particularly the GCC) becoming the target of cyber-criminal activity is the growth of international banking and money laundering [10]. This allows unique opportunities for cybercriminals to exploit the advanced financial infrastructure (which allows the rapid transfer of money to any country) without fear of detection. Electronic transfer is also an efficient tool for hiding the sources of money laundering. The Internet has experienced a large number of cases of money laundering, including victims in the Middle East who is deceived in order for their identity to be stolen, or for money to be transferred from legitimate accounts using phishing and fraud. For example, the attacker sends a link specifically aimed at users in the Middle East, accompanied by an email supposedly asking for help to transfer money into a U.A.E. bank account. In many cases, users will respond, begin to interact with the criminals, and ultimately become a victim of this scam.

A small number of countries in the Middle East, including the U.A.E. and Saudi Arabia, have attempted to form new legislation and legal definitions of crime, however, there is still a need for more specific laws to address the activity of cybercrime. Due to the political issues in the Middle East, many countries in the region are currently employing emergency laws (i.e. against bloggers who are considered to be insulting), rather than laws to address online crimes against ordinary citizens [16]. Other States in the region attempt to prevent such activities by blocking access to some sites.

However, both procedures have proved ineffective, as ordinary laws and emergency regulations in the Middle East have not been designed specifically to deal with cybercrime, and there is no definition of this activity within the law. At the same time, lack of understanding has the potential to criminalize innocent private citizens [17].

B. Cybercrimes inspiration

Terrorism plays a considerable role in Internet crime in the Middle East, both as a means of communication, and as a weapon. Cyber terrorism is growing in the region as a result of religious motives, including those relating to the conflict between Israel and Palestinian, other political issues, along with unemployment. 'Jihad online' claims to use hacking technology to make jihad against their enemies [14]. Cyber terrorists use their websites on the Internet for many activities, including psychological warfare; propaganda; recruitment; fundraising; coordination; and data mining. One Jihad website captures information about users surfing their websites, connecting with those who appear most interested in the group, and therefore likely candidates for recruitment. Internet technology can be used to target recruits through interactive facilities (e.g. chat rooms and Internet cafes), identifying receptive individuals, and in particular young people with a religious motivation, which can potentially be converted into becoming a terrorist. They also look for information technology professionals who could be influenced to assist them with technology. The use of electronic bulletin boards and forums gives openings to potential recruits.

Due to the increasing issues concerning a lack of security awareness and unemployment among users, cybercriminals in the region are constantly seeking new methods of stealing. Since most hackers are 'script kiddies', they use hacking sites to organize crime on the Arabic Internet and disseminate their activities. Spam and trolling activities have now become the largest problem in the region. Local cybercriminals target home users, websites, and the websites of financial institutions and small businesses. Internet crimes for financial gain in the Middle East include violating intellectual property rights through the selling of pirated software. Cybercriminals also use spam to sell banned products, i.e. drugs and pornography. Spam can also be used by cybercriminals to generate traffic, and steal money, from advertisement networks that begin with an awareness of the problem and block Internet traffic from some of the sponsors of the Middle East. The Middle East is also known for credit card fraud and criminals intercepting electronic payment methods to Middle East countries. . According to (Hassan 2015), the number of cybercrime cases in the Middle East is increasing from 2010 to 2014 by more than double [22], as shown in table I.

TABLE I. THE NUMBER OF CYBERCRIME CASES

The number of cybercrime cases		
No.	Year	Number of cases
1	2010	371
2	2011	300
3	2012	563
4	2013	997
5	2014	1212
Total Number		3443

C. Government Sector

Middle Eastern governments incur high costs on an annual basis to restart and repair the breakdown of machinery and interests due to cybercrime [18]. Cybercrime can be directed towards relevant government bodies, in order to generate confusion and attempt to destabilize security and stability, and cause the state financial losses, i.e. access to classified information, deleting, or modifying, them for criminal ends, including: support for terrorism and extremist ideas; spreading rumors; disabling vital systems of government; and disabling and sabotaging servers containing information.

Saudi Arabia and the UAE are first and second, respectively, in the exposure of electronic crimes, at the level of GCC. Trend Micro Inc., a well-known statistical company in the field of protection and the fight against viruses, has indicated the presence of over 700 thousand cases of systematic collapse in Saudi Arabia alone during a period of nine months, thus placing Arabia in first place (up to 64%), followed by the UAE at 20% [19]. In addition, international statistics in 2014 reveal the dangers of electronic piracy on economic institutions, stressing that Egypt's institutions have suffered the highest number of intrusion-mails in the Middle East during 2014. This has alerted the Egyptian government, and some of the regulatory institutions, primarily the Central Bank of Egypt, to develop a strategy for the security of information systems to combat the risk of the penetration of mail [20]. Thus, instability in a country leads to increased cybercrime.

D. Private Sector

Electronic crime and internal threats are a considerable challenge, and only 40% of companies in the Middle East and North Africa have perceived a growth in this danger, compared to 60% at the global level. It may be that the disparity in this ratio is due to a lack of awareness of this threat in the region, in addition to more stringent legislation in other regions of the world. The business sector in the regions of the Middle East and North Africa need to increase awareness of these risks through management, and a willingness to deal with the expected increase of regulatory pressures to be applied in future. The use of data analysis tools needs to become an essential aspect of risk management procedures and compliance programs, to be simultaneously applied proactively and interactively [19]

Events occurring in the Arab region are no less dangerous than those occurring around the world in terms of the Kaspersky Report, which refers to the exposure of banks in Morocco, in particular, to large-scale piracy operations over the

past two years. The report also considers criminal organizations in the Middle East seeking to launch electronic attacks on banks and companies operating in the Arab region in the future [23].

VII. CONCLUSION AND FUTURE WORK

Given the political and economically volatile situation in the Middle East, due to the daily use of a large number of computers and the Internet, the Middle East is vulnerable to considerable numbers of electronic attacks in the future. This is of concern on both an individual, and collective, level, and particularly in light of the weakness of the laws relating to electronic crimes and weak information technologies dedicated to deter these crimes and regulations. It will thus be necessary to focus on the subject of electronic crimes, and the establishment of intensive awareness programs, in addition to strengthening the IT infrastructure in the Middle East, focusing on reducing the anticipated dangers of e-cybercrime.

The methodology for future research is empirical, with data being collected through questionnaires and interviews. The selected population for future study will be the e-government community of the government sector of the Middle East. A Likert scale questionnaire will be used to collect data from the targeted population. Different mediums will be used to distribute a questionnaire in order to ensure the response rate is 100%. It is considered that the factors include (but are not limited to): viruses; malware; spam; hacking; DNS; phishing; spoofing; and cyber terrorism. Two types of questionnaires will be collected. The common type of current and potential cybercrimes will be defined, based on the analyzed data, and the solution framework will be recommended to enhance the security level against cybercrimes.

REFERENCES

- [1] Liang, B. and Lu, H., 2010. Internet development, censorship, and cyber crimes in China. *Journal of Contemporary Criminal Justice*, 26(1), pp.103-120.
- [2] Reddy, G.N. and Reddy, G.J. (2014). A Study of Cyber Security Challenges and Its Emerging Trends on Latest Technologies. *arXiv preprint arXiv:1402.1842*.
- [3] Saini, H., Rao, Y.S. and Panda, T.C. (2012). Cyber-crimes and their impacts: A review. *International Journal of Engineering Research and Applications*, 2, pp.202-209.
- [4] Nain, N. and Batra, S. (2013). Preliminary Study of Cyber Crime, its Impact and Remedial Acts in India. *International journal of computer science and communication engineering, special issue*, pp.223-227
- [5] Aggarwal, P., Arora, P. and Ghai, R. (2014). Review on Cyber Crime and Security. *International Journal of Research in Engineering and Applied Sciences*, 2(1), pp.48-51.
- [6] Jeong, C.H. (2007). *Fundamental of development administration*
- [7] Bell, D. (2004). *Cyberculture: The key concepts*. Psychology Press
- [8] EVERS, J., January 19, (2006), 2006-last update, Computer crimes cost 67 billion, FBI says [Homepage of Cnet], [Online]. Available: http://news.cnet.com/2100-7349_3-6028946.html [01/31, 2014].
- [9] Morrie, J. (1988). *Building a Secure Computer System(PDF)*. Van Nostrand Reinhold. p. 3. ISBN 0-442-23022-2. Retrieved 6 September 2015
- [10] El-Guindy, M.N. (2008). *Cybercrime in the Middle East*. ISSA J, 17.
- [11] Internet World Stats. (2016). <http://www.internetworldstats.com/stats5.htm>. Accessed [March 2016]
- [12] MORALES, J.A. (2008). A behavior based approach to virus detection, Florida International University.
- [13] COHEN, F. (1987). Computer viruses: theory and experiments. *Computers & Security*, 6 (1), pp. 22-35.
- [14] Jihad Online: Islamic Terrorists and the Internet , <http://archive.adl.org/internet/jihad.html>. Accessed [March 2016]
- [15] World Bank report, 2007 <http://www.worldbank.org>. Accessed [March 2016]
- [16] Egypt blogger jailed for 'insult' http://news.bbc.co.uk/2/hi/middle_east/6385849.stm. Accessed [March 2016]
- [17] Arabic free Internet initiative, <http://openarab.net/>. Accessed [March 2016]
- [18] Mona Shaker (2010), "The impact of electronic crime on the economic aspects", Center of Excellence in Information Assurance. <http://goo.gl/indM7k>. Accessed [March 2016]
- [19] Electronic intrusions and internal threats (2016), <http://www.alriyadh.com/1126124>. Accessed [Feb 2016]
- [20] Egypt leader Middle East countries prone to electronic piracy during 2014 (2014), <http://www.vetogate.com/1583751>. Accessed [March 2016].
- [21] DWYER, P.,2010. CYBER CRIME IN THE MIDDLE EAST.
- [22] Hassan, A. 2015, 1000 victim because of cyber crimes. http://m.aljarida.com/pages/news_more/2012725434/3. Accessed [May 2015].
- [23] Insurance companies are facing a difficult test with the growing piracy losses, <http://goo.gl/buF6yF>. Accessed [March 2016].

Performance Comparison between Forward and Backward Chaining Rule Based Expert System Approaches Over Global Stock Exchanges

Sachin Kamley
Deptt. of Computer Application's
S.A.T.I.,
Vidisha, India

Shailesh Jaloree
Deptt. of Appl. Math's and CS
S.A.T.I.,
Vidisha, India

R.S. Thakur
Deptt. of Computer Application's
M.A.N.I.T.,
Bhopal, India

Abstract— For the last couple of decade's stock market has been considered as a most noticeable research area everywhere throughout the world because of the quickly developing of the economy. Throughout the years, a large portion of the researchers and business analysts have been contributed around there. Extraordinarily, Artificial Intelligence (AI) is the principle overwhelming area of this field. In AI, an expert system is one of the understood and prevalent techniques that copy the human abilities in order to take care of particular issues. In this research study, forward and backward chaining two primary expert system inference methodologies is proposed to stock market issue and Common LISP 3.0 based editors are used for designing an expert system shell. Furthermore, expert systems are tested on four noteworthy global stock exchanges, for example, India, China, Japan and United States (US). In addition, different financial components, for example, Gross Domestic Product (GDP), Unemployment Rate, Inflation Rate and Interest Rate are also considered to build the expert knowledge base system. Finally, experimental results demonstrate that the backward chaining approach has preferable execution performance over forward chaining approach.

Keywords— Stock Market; Artificial Intelligence; Expert System; Macroeconomic Factors; Forward Chaining; Backward Chaining; Common LISP 3.0.

I. INTRODUCTION

In nowadays stock price prediction has become sizzling topic in the time series analysis and dependably stays in the limelight because of rising and falling states of the economy. So different researchers and business experts have paid consideration to break down and anticipating the future estimation of stock exchange prices [1] [2]. Previously, the various tools and techniques have been proposed for the numeric stock value prediction. However, an Artificial Neural Network (ANN) is one of the most prominent techniques among them. For most recent five decades an expert system has risen as effective AI techniques and had demonstrated to its value in different areas such as designing, account, farming, medicine, crystal gazing and numerous more [3].

An expert system is likewise called knowledge based system which uses knowledge to tackle problems and knowledge must be encoded in some forms of facts, rules, procedures, relations, etc. In expert system knowledge might be gathered from different sources, for example, primary

source human expert and secondary sources, for example, books, magazines, newspapers, reputed journals [4]. After knowledge acquisition, its representation is essential and very challenging issues. For symbolic representation of knowledge, there are various techniques are utilized such as frames, semantic net, scripts, production rule and so on. Figure 1 shows the basic structure of expert systems.

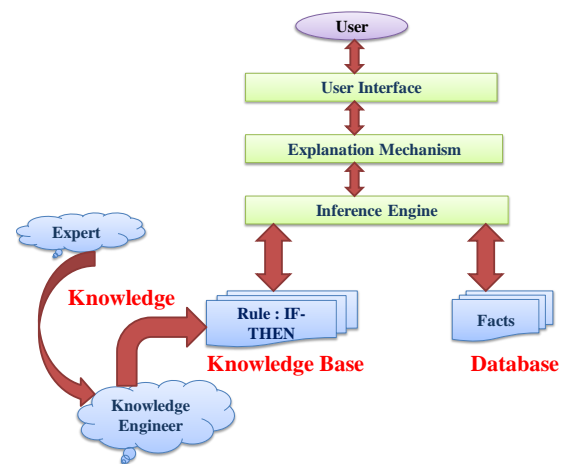


Fig. 1. Basic Structure of an Expert System [3] [5]

In Figure 1 inference engine is the most capable component of the expert system that is in charge of completing the reasoning procedure where the expert system comes to end at specific solution. Additionally, the inference engine is likewise responsible for links the rules with the knowledge base where facts are given by the database.

Throughout the previous couple of years, stock market plays an important role in the quickest developing economy, but accurately shares price forecasting are still very chaotic and complicated process [6]. There are various algorithms are developed for numerical forecast, but no authors have endeavored symbolic representation of stock market data. There are two basic approaches forward chaining and backward chaining are utilized as a part of the expert system design and development.

Forward chaining is a data driven methodology which begins from the known data and continues forward with that

data. The procedure is stopped when no more rules are remaining left to fire. In the forward chaining procedure, firstly premises are specified after that conclusion [3] [7]. The forward chaining rule fashion is given below.

```
IF
  A
  AND
  B
THEN
  C
```

In other words, backward chaining is the goal-driven reasoning where a system begins with the predetermined goal i.e. hypothetical solution and the inference engine attempts to find evidence to support it. In backward chaining procedure, firstly, goal is specified after that premise is proven to support that goal [3] [5] [7]. The backward chaining rule fashion is given below.

```
      C
    IF
      A
      AND
      B
.
```

There are some important considerations for this research study is:

- 1) Various fundamental and technical variables as well as some important macroeconomic factors are also considered to construct the stock market knowledge base.
- 2) Very few authors have attempted a comparative study between forward and backward chaining for the stock market problem.
- 3) The ultimate focus of this research study is to selection of appropriate reasoning strategy for the stock market as well as improving the inference capability of a stock market expert system.
- 4) The global stock exchanges, for example, India, China, Japan and US are selected for study purposes.
- 5) This research study will be contributed to already ongoing research on the stock market field.
- 6) This research study will provide a foundation for novice stock users and researchers by providing a clear picture of expert system tools and techniques.

II. LITERATURE REVIEW

In this section we are presenting some significant researchers work over the years.

Niederlinski (2001) [8] has presented forward and backward reasoning approaches for expert system shell. The 24 production rules with uncertainty factor are constructed to draw expert knowledge base. Therefore, modified standard certainty factor algebra is used for modelling the knowledge

base. However, the most important property of the system is that it might be capable to automatically draw conclusions from the knowledge base. Moreover, the report generation facility is also incorporated in the system for further future analysis.

Erdani (2012) [9] has proposed backward chaining algorithm for the ternary grid expert system. He used the recursive procedure to design the backward chaining inference engine. The experimental results showed that his system can efficiently work with ternary grid knowledge base and inference process works better than previous algorithms.

Zarandi et al. (2012) [10] have presented fuzzy expert system shell for evaluating the intellectual capital. They used various fuzzy linguistic variables to express the level of qualitative evaluation and expert criterion. The knowledge base is constructed with various factors such as capital structure, market share, rate, investors, knowledge, customer capital, etc. the experimental results showed that system had better performance in terms of linguistic variables and system has also extended power over the linguistic variables for future study.

Ajlan (2015) [11] has presented the comparative study between forward and backward chaining for education expert system. The various academic performance indicators such as grade point, attendance, course, etc. are used by him to construct the knowledge base. Finally, experimental results showed that forward reasoning strategy is more appropriate than backward reasoning in terms of deriving goals.

Kamley et al. (2015) [12] have presented forward chaining rule based expert system approach for Bombay Stock Exchange (BSE) of India index. The knowledge base consists of approximately ten rules and rules are comprised of basic stock variables such as open price, close price, high price, low price and volume.

Kamley et al. (2015) [13] have presented a comparative study between forward and backward chaining strategies over National Stock Exchange (NSE) of India forecasting and knowledge representation. In their study, they incorporate the various fundamental factors such as inflation rate, oil prices, earnings per share, dividend and US dollar prices. The knowledge base is constructed with these factors. The common LISP 3.0 editor is employed for expert system design task. The experimental results showed that backward chaining strategy is more appropriate than forward chaining strategy in terms of number of iterations.

The section 3 describes the knowledge base construction section 4 describes proposed methodologies with examples, section 5 describes experimental results and at last section 6 describes conclusion and future scopes of study.

III. KNOWLEDGE BASE CONSTRUCTION

A quality of expert system depends on its knowledge base and quality of knowledge base depends on knowledge which

is fed by knowledge engineers. The knowledge base is maintained using a set of rules and using a variable name list [14]. Table I shows the stock market variable description.

Table I. Stock Market Variable Description

S.No.	Variable	Description
1	OP	Open Price
2	HP	High Price
3	LP	Low Price
4	CP	Close Price
5	VOL	Volume
6	INFL	Inflation
7	INTRST	Interest Rates
8	IMP	Imports
9	EXP	Exports
10	GDP	Gross Domestic Product
11	OLP	Oil Prices
12	EPS	Earnings Per Share
13	DIVD	Dividend
14	USDP	US Dollar Prices
15	STK	Stock Market
16	FDI	Foreign Direct Investment
17	UNEMPR	Unemployment Rate
18	ECOMY	Economy

After variable description, stock market knowledge base contains a set of rules. Stock market domain is unpredictable and consisting of so many technical variables and rules so representation and mapping of all these rules are not possible. In this study, 50 sample rules are considered for stock market knowledge base [15] [16]. Table II shows a sample of stock market knowledge base.

Table II. Sample of Stock Market Knowledge Base

Rule No.	Rule Description
1	IF IMP fall and EXP fall THEN STK down
2	IF INFL fall and USDP rise THEN STK UP
3	IF INFL fall and USDP fall THEN STK down
4	IF IMP fall THEN STK down
5	IF USDP rise THEN STK up
6	IF ECOMY fall and USDP rise THEN STK up
7	IF INTRST rise and INFL rise THEN STK down
8	IF IMP fall and ECOMY fall THEN STK down
9	IF OP rise THEN share prices rise
10	IF OP and LP and CP rise THEN sell shares of all indices
13	IF EPS and DIVD fall THEN fall shares of all indices
14	IF VOL and OP rise THEN shares rise of all indices
15	IF OLP rise THEN fall shares of all indices

IV. PROPOSED METHODOLOGIES

Presently days, stock market data has grown to a larger extent. So therefore, need a sort of inference procedures and formal methods which can be searched very fast and infer results in less time. In this study, forward and backward chaining two mainstream expert system inference procedures are adopted. This section clearly describes both approaches in detail manner.

The Forward Chaining procedure begins by applying the rules in a forward direction. It is a top down procedure, i.e. recursively applying the rules over data to generate more data. Algorithm 1 shows forward chaining inference procedure [3] [17].

Algorithm 1. Forward Chaining Inference Procedure

- 1) Start
- 2) Check the premises of each rule against the knowledge base whose IF part is satisfied with the current contents of the fact base.
- 3) Conflict resolution occurs if more than one rule is applicable then rules are fired based on following criteria:
 - 3.1) don't fire a rule twice on the same data.
 - 3.2) fire rules on more recent working memory elements before older ones.
 - 3.3) more specific preconditions rules fires before ones with more general preconditions.
- 4) Execute the rule and finally add the facts in the fact base i.e. facts which are specified in the THEN parts of the conclusion.
- 5) Stop the procedure if no more rules are applicable.

Figure 2 shows a flowchart of a forward chaining procedure.

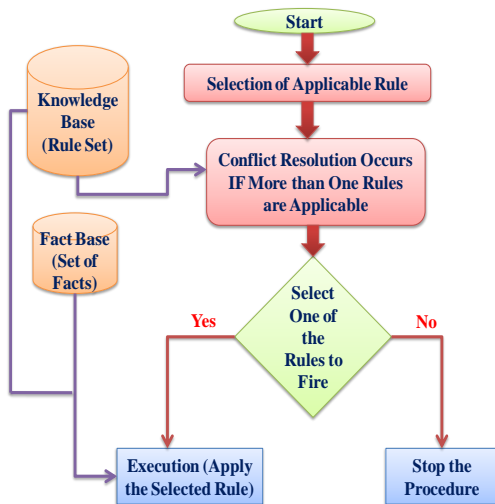


Fig. 2. Flow Chart of Forward Chaining Inference Procedure

Now let's understand the forward chaining procedure with Example 1.

Example 1:

Rule 1: IF P and Q → S

Rule 2: IF P and R → C

Rule 3: IF M → R

Rule 4: IF C → N

Prove that "IF P and M True THEN N is True"

Solution:

1st Iteration:

- 1) Initially, the only input P and M is true.
- 2) Start with Rule 1 and go to forward direction until Rule "fires" is found.
- 3) Rule 3 fires and conclusion R are found.
- 4) No other rules fired and iteration one ends here.

2nd Iteration:

- 1) Rule 2 is fired and conclusion C is found.
- 2) Rule 4 fires and conclusion N are found.
- 3) The procedure stops because goal N is true.

The backward Chaining procedure begins by applying the rules in a backward direction, i.e. taking a goal or query and recursively working backwards to find more evidences in order to satisfy the goal or hypothesis. In other words, backward chaining procedure works from the goal back to facts. The Algorithm 2 describes the backward chaining inference procedure [5] [18] [19].

Algorithm 2. Backward Chaining Inference Procedure

- 1) Start with the top goal on the stack, i.e. checks the conclusions of the all rules those make to satisfy the top goal on the stack.
- 2) Do the process to check out each rule one at a time.
 - 2.1) evaluate the conditions of the each rule of the IF part (antecedent) one at a time.
 - 2.2) currently IF condition is unknown, i.e. there is no sufficient information available to determine whether the condition is to be satisfied. Now push a goal on the goal stack in order to make that condition known to be true and recursively invoke the procedure whether information is available to satisfy the goal.
 - 2.3) if the condition is not satisfied, then go to step 2 and repeat procedure from beginning.
 - 2.4) if the condition is satisfied for applicable rules then add facts in the working memory (i.e. facts specified in the THEN parts of the rule).
- 3) Pop the top goal from the stack and return back from the procedure.
- 4) Finally system flashes "success" message and goal stack will be empty.

Figure 3 shows a flowchart of backward chaining inference procedure.

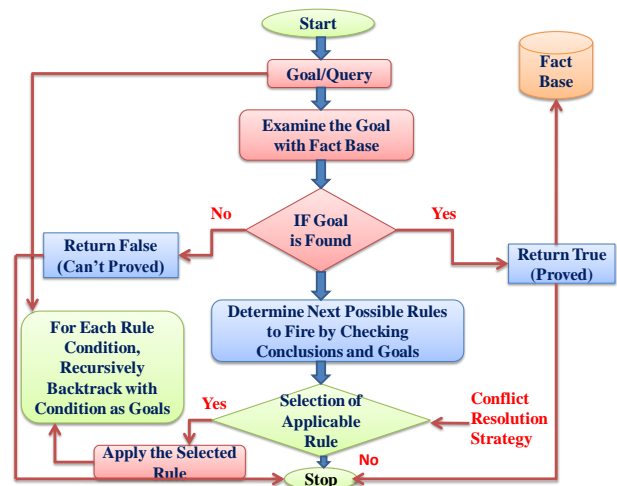


Fig. 3. Flowchart of Backward Chaining Inference Procedure

Now let's understand the backward chaining procedure by same example.

Solution:

1st Iteration:

- 1) Start with goal N is true and go backward until rule "fires" is found.
- 2) Rule 4 is fired and one new sub goal C is added to prove. Go backward and search the rules.

- 3) Rule 2 is fired and conclusion P is true. New sub goal R is added in the working memory to prove it. Go backward and search the rule.
- 4) No other rules fired and iteration one ends here.

2nd Iteration:

- 1) Rule 3 is fired and conclusion M is becoming true (2nd input found here).
- 2) Both inputs P and M are ascertained.
- 3) The procedure stops because goal N is proved.

V. EXPERIMENTAL RESULTS

In this study, stock market data are utilized for mapping and representation of knowledge base. Throughout the years, List Processing (LISP) is the general purpose and symbol oriented language which has been being used extensively for the knowledge representation task. However, the Common LISP 3.0 based editor is utilized for stock market expert system design and development task [4] [20] [21]. All the input facts are stored in fact base. The Figure 4 shows sample of fact base generated from LISP environment.

```
Common Lisp-[Lisp Worksheet]
;; Corman Lisp 3.01 (Patch level 0)
;; Copyright © Corman Technologies Inc. All rights reserved.
;; Unlicensed version, evaluation period expires in 24 days.

(defvar *fact-list *)
(defvar *rule-list *)
(setq *fact-list * '(
(is fall (close prices, high prices))
(is rise (inflation rates, share prices))
(is fall (inflation rate, share prices))
(is rise (volume, open prices))
(is rise (open prices))
(is rise ( (volume), (share open prices))
(is fall ( (volume), (share close prices))
(is rise (US dollar prices, interest rate))
(is fall (oil prices))
.....
.....
.....
))
```

Fig. 4. Sample of Fact Base Generation

The above mentioned rules are inputted into stock knowledge base in the rule base format. The inference engine is responsible for searching all the applicable rules in the knowledge base and finally updated the fact base after execution of rules. Figure 5 shows the rule base (knowledge base) representation from LISP environment.

```
Common Lisp-[Lisp Worksheet]
;; Corman Lisp 3.01 (Patch level 0)
;; Copyright © Corman Technologies Inc. All rights reserved.
;; Unlicensed version, evaluation period expires in 24 days.

(setq *rule-list* '(
(R1 IF (is rise INFL THEN (is fall share prices))
(R2 IF (is fall INFL THEN (is rise share prices))
(R3 IF (is rise share OP) THEN (is sell shares of all indices))
(R4 IF (is rise (OP LP) and (is rise CP) THEN (is buy shares of all indices))
(R5 IF (is rise (VOL OP ) THEN ((is sell shares) and (is earn profit)))
(R6 IF (is fall (USDP) THEN (is rise share prices))
(R7 IF (is rise (USDP ) THEN (is fall share prices))
(R8 IF (is rise (OLP) THEN (is fall share prices))
(R9 IF (is fall (OLP) THEN (is rise share prices))
(R10 IF (is rise EPS DIVD) THEN (is rise share prices))
.....
.....
))
```

Fig. 5. Sample of Rule Base Generation

After a generation of fact base and rule base users will be interacting with system and ask various questions about the system. The questions asked by the system are “what will be effect of falling US Dollar prices”, “what will be impact on India of rising open and high prices of NASDAQ and Hang Sang index”, “what will be effect of rising oil prices on BSE index”, “what will be effect of falling volume and close prices”, “what will be impact on all index of falling inflation rate and interest rate”, “what will be effect of rising of GDP”, “what will be effect of rising of unemployment rate on all indices”, “what will be effect of rising of FDI on all indices” [22] [23]. The expert system tries to solve these queries one by one and inference engine procedure searches all the rules in the knowledge base. Ultimately, all applicable rules are to be executed and respective conclusion is drawn and fact base is updated with the new knowledge. Figure 6 shows execution of forward chaining inference procedure.

```
Common -LISP
;; Corman Lisp 3.01 (Patch level 0)
;; Copyright © Corman Technologies Inc. All rights reserved.
;; Unlicensed version, evaluation period expires in 24 days.
;;Could not load auto-update index file. This may be because you are not connected to the internet, or because you need to configure proxy server settings

->(procedure FC_Infer (rule-base, fact-base)
rules ← Select (rule-base, fact-base);
while rules ≠ ∅ do
rule ← resolveconflicts(rules);
Apply(rule);
rules ← Select(rule-base, fact-base)
od
end
->(setq *FC_Infer *T); // procedure is invoked when command is entered
```

Fig. 6. Forward Chaining Inference Procedure

In Figure 6 inference engine procedure starts by searching inference rules until it finds one where the IF clause is known to be true. When found it can conclude, or infer the THEN

clause resulting in the addition of new information to the fact base. Figure 7 shows initial production system environment.

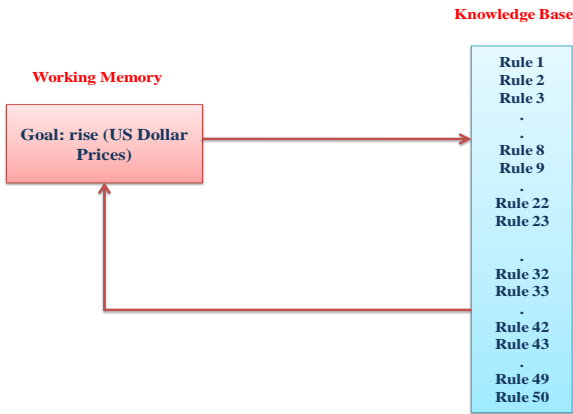


Fig.7. Production System Environment

After production system environment, the inference process searches specified goal sequentially in the knowledge base. As a result, 6 rules fired and new information is to be added in the working memory. The Figure 8 shows production system environment after rule invocation.

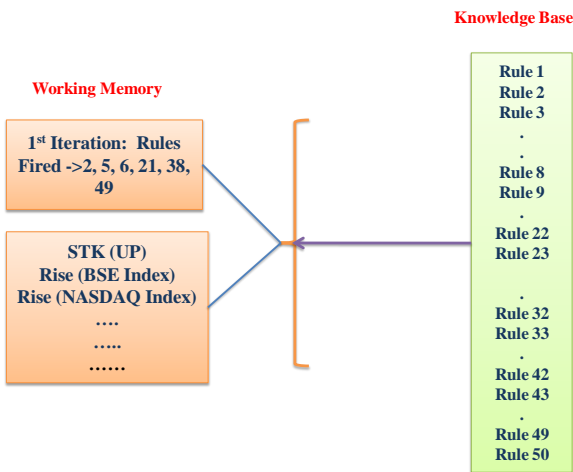


Fig. 8. Production System Environment after Rule Invocation

Therefore, inference process continues until goal does not become true and all possible facts have been asked to be fired. In this manner, all goals are attempted. Table III shows descriptive statistics of forward chaining inference procedure.

Table III. Descriptive Statistics of Forward Chaining Inference Procedure

S.No.	No. of Fact	No. of Rule	No. of Iteration
1	20	10	30
2	35	20	55
3	70	30	82
4	120	50	130

Figure 9 shows a bar graph for rule no. against no. of iterations.



Fig. 9. A Bar Graph for Rule No. Against No. of Iterations

Figure 9 states that when no. of rules increase then no. of iterations also increase. Figure 10 shows the invocation of backward inference procedure.

```
Common Lisp-[Lisp Worksheet]
;; Corman Lisp 3.01 (Patch level 0)
;; Copyright © Corman Technologies Inc. All rights reserved.
;; Unlicensed version, evaluation period expires in 24 days.
;; Could not load auto-update index file. This may be because you are not connected
;; to the internet, or because you need to configure proxy server settings
-> (Defun backward_debug ()
  (IF (execute-rule rule-list)
    (backward)))
(defun execute-rule ()
  (find-if #'eval-rule-f "rule-list *))
-> (setq *backward_debug* T); //procedure is
automatically invoked when command is entered
```

Fig. 10. Invocation of Backward Inference Procedure

In Figure 10 backward chaining procedure begins by searching the inference rules until it finds one which has a THEN clause that matches a desired goal. If the (IF clause) of that inference rule is not known to be true, then it is added to the list of sub goals. Figure 11 shows initial backward chaining production system environment.

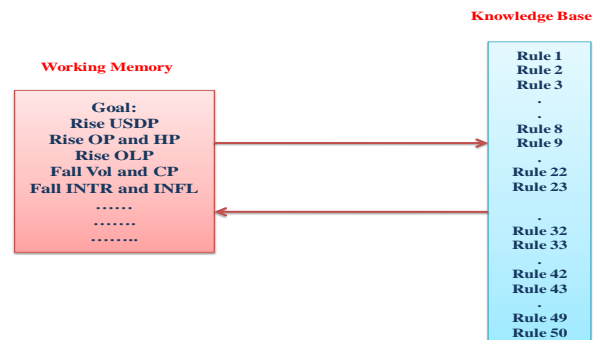


Fig. 11. Backward Chaining Production System Environment

Afterwards, production system environment the inference procedure searches the specified goal in the knowledge base. As a result, 12 rules are fired and new sub goals are to be added in the working memory. Figure 12 shows backward production system after rule invocation.

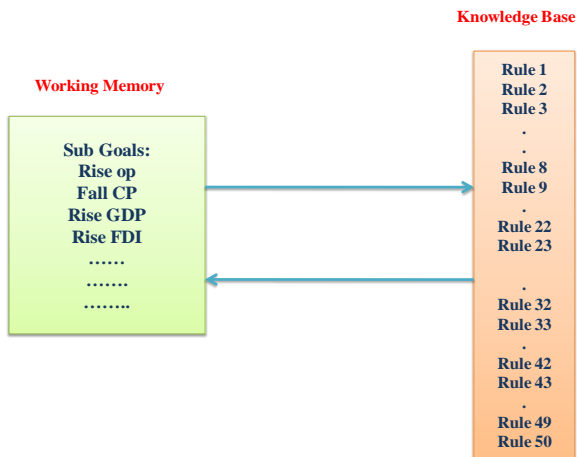


Fig. 12. Backward Production System after Rules Invocation

Therefore, the procedure continues until all sub goals have been tested. At the end procedure terminates and all goals are attempted. Table IV shows descriptive statistics for the backward chaining inference procedure.

Table IV. Descriptive Statistics of Backward Chaining Inference Procedure

S.No.	No. of Fact	No. of Rule	No. of Iteration
1	5	10	15
2	10	20	25
3	20	30	45
4	35	50	85

Figure 13 shows a bar graph for rule no. against no. of iterations.

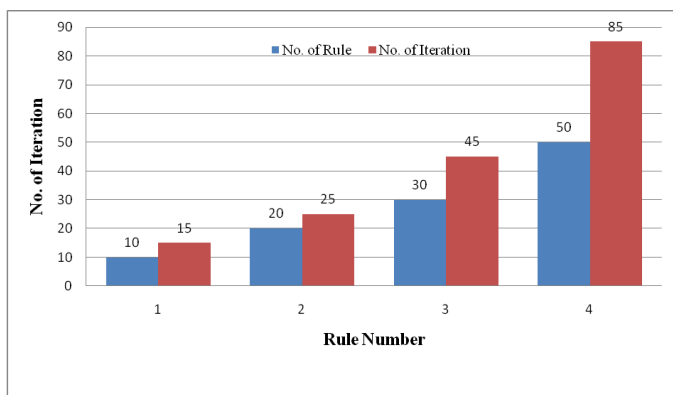


Fig. 13. A Bar Graph for Rule No. Against No. of Iterations

Figure 13 states the effect of increasing the rule on the no. of iterations i.e. if no. of rules increase than no. of iterations also increases. Figure 14 shows the performance comparison between forward and backward chaining approach.

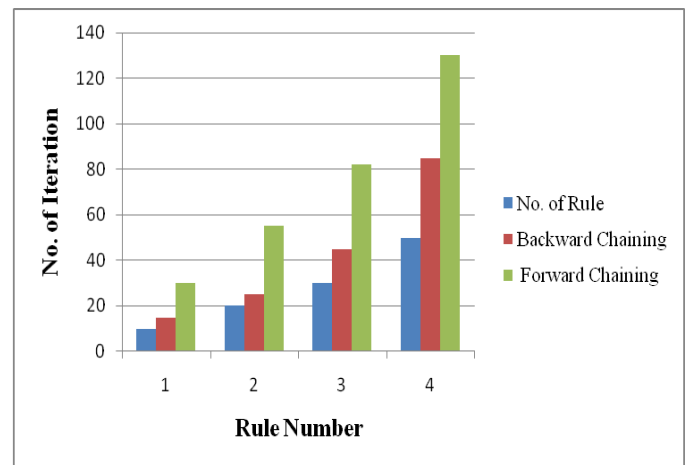


Fig. 14. Performance Comparison between Forward and Backward Chaining Process

Figure 14 depicts that the backward chaining method performs reasonably well in comparison to the forward chaining method, i.e. when no of rules increase than no. of iterations decrease. Table V shows forecasted results for stock market expert system.

Table V. Forecasted Results for Stock Market Expert System

S.No	Fact	Inference Result	Investment Decision
1	Rise USDP	Fall BSE Index	Buy
2	Rise OP and HP	STK UP and rising share prices	Sell
3	Rise OLP	STK Down	Buy
4	Fall VOL and CP	STK DOWN	Buy
5	Fall INFL and INTR	STK UP	Sell
6	Rise GDP	Rise BSE Index	Sell
7	Rise UNEMPR	Fall BSE Index	Buy
8	Rise FDI	Rise BSE Index	Sell
9	Fall OP and CP	Fall BSE Index	Buy
10	Fall LP and Vol	Fall BSE Index	Buy
11	Rise USDP and INFL	Fall Hang Sang Index	Hold
12	Fall GDP and FDI	Fall Nikkei 225 Index	Hold
13	Rise GDP and USDP	Rise NASDAQ Index	Sell
14	Fall FDI and GDP	STK DOWN	Buy
15	Fall INFL	STK UP	Sell

Table V shows the major performances of expert system and also shows the best opportunity for stock market users to invest money in the near future.

VI. CONCLUSION AND FUTURE SCOPES

In this study four noteworthy world stock exchange data are considered for modeling the stock market knowledge base.

The LISP based editor is utilized to model the forward and backward chaining approaches. Based on the findings, it is found that backward chaining strategy has better execution performance in contrast with forward chaining strategy. The only disadvantage of forward chaining strategy is found that it triggers the rules sequentially and takes more time to infer the goal while the backward chaining strategy searches the desired goal in an extremely fast i.e. takes less time to infer the goal. Besides, this research study will set out the establishment for all stock investors and brokers and also will be useful to invest money in the share market. In the future, experimental results will be improved by utilizing more fundamental and technical factors and in addition suitable knowledge representation techniques like frame based approach will be considered for expert system design task.

REFERENCES

- [1] Brown and Jennings, "On Technical Analysis", *The Review of Financial Studies*, Vol. 2, Issue (4), 1989, pp. 527-551.
- [2] L. Dymova, P. Sevastianov and K. Kaczmarek, "A Stock Trading Expert System Based on the Rule-Base Evidential Reasoning Using Level 2 Quotes", *Expert System with Applications: An International Journal*, Vol. 39, Issue (8), June 2012.
- [3] E. Rich and K. Knight, "Artificial Intelligence", 3rd Edition, *McGraw-Hill*, 1991.
- [4] S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", 3rd Edition, *Prentice Hall*, 2009.
- [5] J. Nils. Nilsson, "Artificial Intelligence A New Synthesis", 2nd Edition, *Morgan Kaufmann Publishers*, 1998.
- [6] G. C. Yunusoglu and H. Selim, "A Fuzzy Rule Based Expert System for Stock Evaluation and Portfolio Construction: An Application to Istanbul Stock Exchange", *Expert System with Applications*, Vol. 40, Issue (3), pp. 908-920, 2013.
- [7] P. Tanwar, T.V. Prasad, M.S. Aswal, "Comparative Study of Three Declarative Knowledge Representation Techniques", *International Journal on Computer Science and Engineering (IJCSSE)*, Vol. 2, No.7, pp. 2274-2281, 2010.
- [8] A. Niederlinski, "An Expert System Shell for Uncertain Rule and Model Based Reasoning", *Methods of Artificial Intelligence in Mechanics and Mechanical Engineering*, Vol. 8, pp. 1-10, 2001.
- [9] Y. Erdani, "Developing Backward Chaining Algorithm of Inference Engine in Ternary Grid Expert System", *International Journal of Advanced Computer Science and Applications*, Vol. 3, No. 9, pp. 241-245, 2012.
- [10] F.M. Zarandi, M.H. Neda, S. Bastani, "A Fuzzy Rule Based Expert System for Evaluating Intellectual Capital", *Advances in Fuzzy System*, Vol. 12, pp. 1-11, 2012, doi:10.1155/2012/823052.
- [11] A. Ajlan, "The Comparison between Forward and Backward Chaining", *International Journal of Machine Learning and Computing*, Vol. 5, No. 2, pp. 106-113, 2015.
- [12] S. Kamley, R.S. Thakur, S. Jaloree, "Rule Based Approach for Stock Selection: An Expert System", *International Journal of Computing Algorithm (IJCOA)*, Special Issue, Vol. 4, pp. 1142-1146, 2015.
- [13] S. Kamley, S. Jaloree, K. Saxena, R.S. Thakur, "Forward Chaining and Backward Chaining: Rule Based Expert System Approaches for Share Forecasting and Knowledge Representation", *Presented for 7th International Conference on Quality, Reliability, Infocom Technology and Business Operations (ICQRIT)*, Sponsored by Springer, University of Delhi, 28-30 Dec, 2015.
- [14] A.P. Das, "Security analysis and portfolio Management", *I.K. International Publication*, 3rd Edition, 2008.
- [15] Online Stock Market Dataset Available on Yahoo Finance Site, "http://www.yahoofinance.com", Accessed on 2/01/2016.
- [16] Online Macroeconomic Variables Data Available on Site, "http://www.indexmondi.com", Accessed on 5/01/2016.
- [17] S.M. Aqil Burney and N. Mahmood, "A Brief History of Mathematical Logic and Applications of Logic in CS/IT", *Journal of Science*, Vol. 34, Issue (1), pp. 61-75, July, 2006.
- [18] S. Ali and L. Iwanska, "Knowledge Representation for Natural Language Processing in Implemented System", *Natural Language Engineering*, 3:97-101, Cambridge University Press, 1997.
- [19] H. Shi, K. Meli, and A. Steven, "Scalable Backward Chaining Based Reasoner for a Semantic Web", *International Journal on Advances in Intelligent Systems*, Vol. 7, pp. 23-38, 2014.
- [20] Common Lisp Compiler Downloaded from "http://www.commonlisp.com", Accessed on 25/10/2015.
- [21] H. Boley, "Expert System Shells: Very High Level Languages for Artificial Intelligence", *Expert System*, Vol. 7, No. 1, pp. 2-8, Feb. 1990.
- [22] P.P. Bonissone and R.M. Tong, "Reasoning with Uncertainty in Expert Systems", *International Journal of Man-Machine Studies*, Vol. 22, No. 3, pp. 241-250, 1985.
- [23] S. Gryglewicz, "A Theory of Corporate Financial Decisions with Liquidity and Solvency Concerns", *Journal of Financial Economics*, Vol. 99, Issue (2), pp. 365-384, 2011.

AUTHORS PROFILE



Sachin Kamley did his Masters from S.A.T.I., Computer Applications Department, Rajiv Gandhi Technological University, Bhopal (M.P.) in 2006. He is working at Samrat Ashok Technological Institute (S.A.T.I), Vidisha as a Lecturer from May 2007 to Department of Computer Applications and also completed the Ph.D. from Barkatullah University, Bhopal in the year 2015 to till date. He has attended many workshops and conferences of

National repute.



Shailesh Jaloree is an Associate Professor in the Department of Applied Math's and Computer Science at Samrat Ashok Technological Institute (S.A.T.I), Vidisha, India. He earned his Master Degree from Devi Ahiliya University, Indore (M.P.) in 1991 and the Ph.D. Degree (Applied Maths) From Barkatullah University, Bhopal (M.P.) in 2002. At Present he is guiding several Ph.D. Research Scholars in Mathematics and Computer Science field. He has published more than 35 Research

Paper in National, International, Journals and Conferences. His areas of interest include Special Function, Data Mining, Data Warehousing and Web Mining.



Ramjeevan Singh Thakur is an Associate Professor in the Department of Computer Applications at Maulana Azad National Institute of Technology, Bhopal, India. He had a long career in teaching and research, including Three Year Teaching in the Department of Computer Applications at National Institute of Technology, Tiruchirappalli, Tamilnadu, India. At Present he is guiding several Ph.D. Research Scholars and handling Government Research Projects of about

Rs. One Crore. He has published more than 75 Research Paper in National, International, Journals and Conferences. He has visited several Universities in USA, Hong Kong, Iran, Thailand, Malaysia, and Singapore.

Analysis of Impact of Varying CBR Traffic with OLSR & ZRP

Rakhi Purohit

Research Scholar,
Department of Computer Science & Engineering,
Suresh Gyan Vihar University,
Jaipur, Rajasthan, India

Bright Keswani, Ph.D.

Associate Professor & Head
Department of Computer Application
Suresh Gyan Vihar University
Jaipur, Rajasthan, India

Abstract— Mobile ad hoc network is the way to interconnect various independent nodes. This network is decentralize and not follows any fixed infrastructure. All the routing functionality are controlled by all the nodes. Here nodes can be volatile in nature so they can change place in network and effect network architecture. Routing in mobile ad hoc network is very much dependent on its protocols which can be proactive and reactive as well as with both features. This work consist of analysis of protocols have analyzed in different scenarios with varying data traffic in the network. Here OLSR protocol has taken as proactive and ZRP as Hybrid protocol. Some of the calculation metrics have evaluated for this analysis. This analysis has performed on well-known network simulator NS2.

Index Terms:- Mobile ad hoc network, Routing, OLSR, Simulation, and NS2.

I. INTRODUCTION

In the current scenario of wireless networking technology where increase of use of personal digital assistant devices like laptops, tablets, mobile phones etc. These devices need the networking on ad hoc basis so here ad hoc networking techniques are very popular. Every node in this type of network is like a router and can work like a relay station which can initiate the transmission as well as can receiver for other node who transmit message [1]. In this type of network nodes can be volatile in nature so they can change their location and its effects on infrastructure of network which frequently change so connection in between nodes are frequently breaks and reform on requirement. This way of networking is very much helpful in some crucial situation like disaster, crowd, battle field etc.

To maintain connectivity and routing in this network several rules have proposed which known as routing protocols which can specify in various ways of their working techniques. Basically these protocols have two types of features, they can be proactive or they can be reactive[2], but in some cases they perform both feature and known as hybrid in nature.

A. Proactive Protocol –

This type of protocol have feature to be provisionally active and updated for any possible route for required transmission. These protocols frequently maintained updated list of possible destination and their routes details on periodic basis, that is known as routing table so that it is also known as table driven protocol[3]. Example- DSDV, OLSR.

B. Reactive Protocol-

This type of protocols reacts according to need of transmission and initiate discovery of route for destination so that it is known as reactive protocol. It initiate work on demand of transmission so it also known as on demand protocols. These protocols do not maintain any updated list of possible routes for all destinations so they need to discover route for each transmission [3]. Examples- AODV, DSR.

C. Hybrid Protocol-

This type of protocols has capability of both type of reactive as well as proactive due to its network infrastructure. This type of protocols maintains updated list of possible routes for all destination in current zone and need to initiate route discovery process for the destination outer then current zone [3].

Here two protocols have taken one is proactive in nature that is OLSR and other is hybrid in nature that is ZRP.

II. OLSR

Optimized link state routing OLSR is a good examples of proactive approach [2] of routing in mobile ad hoc networking. OLSR has feature of stability of link state routing and some enhanced feature like immediate route availability on requirement of transmission in network [11]. The OLSR has feature which can minimize the overhead of network by focus on selected nodes for flooding of the traffic in transmission in network. These selected nodes are known as multi-point-relay or MPR which are responsible to retransmit this message to control. Whenever any node receive this message, it determine the need of retransmit according to route and select one to the MRP available in their neighbor list and retransmit only when this message is not belongs to that particular node.

III. OLSR ARCHITECTURE

OLSR Protocol utilizes two types of messages for communication [11]. The TC means topology control message is responsible for maintain updated details of routes in network for all possible destinations in network and the hello which is first type of message used to maintain identity of neighbors updated details. OLSR protocol is so much efficient in the manner of traffic control. It works very efficiently

where network is very dense and MRP approach work well for efficient routing with less network-overhead.

IV. ZRP PROTOCOL

Zone routing protocol is a hybrid routing protocol [2][3]. It has combination of feature of proactive and reactive both type of routing. So it is known as hybrid routing protocol. ZRP [12] utilize the technique of minimize the control overhead on network by proactive routing technique. It has good capability to decrease the network latency which is created by discovery of routes in reactive routing technique. For this feature the ZRP protocol utilizes the way to define zones in network around each node consist of its specified number of neighbors having similar node distance from center node of zone.

ZRP protocol is consisting of some of the subordinate protocol like NDP, IARP, IERP and BRP. In which there are intra-zone and inter-zone routing protocol perform according to its working mechanism by using zones in network. Intra zone routing protocol also known as IARP is working in a specific zone in network and it is of proactive in nature which maintains updated details of all possible destination and their routes inside a particular network zone with the help of Neighbor Discovery Protocol NDP. Inter zone routing protocol also known as IERP is working between various zones in network. This protocol is basically reactive in nature, it initiate route discovery on the need of message transmission, when a node need to transmit message to a particular node which is not present in current zone and it can be outside of current zone then intra zone routing protocol transmit message to nodes present on the border of current zone which is handled by third protocol known as Border cast Resolution Protocol BRP [12]. So the technique of ZRP is depends on some of the subordinate protocols NDP, IARP, IERP and BRP, which together makes it very efficient in all manner.

V. ZRP ARCHITECTURE

ZRP protocol identifies a network as a collection of various zones. Each node is surrounded by a specific zone with a particular radius that specifies the node distance from central node of zone and also specifies the nodes on the border of zone. This zone radius is not measurement of distance between nodes but it is the count of nodes from center node to the border node of a specific zone in network. Some time it happens that one zone overlap to another various zone due to each node can have its own zone. In every zone there can be two type of neighbor nodes interior node and peripheral node or border nodes.

ZRP protocol is collection of several protocol in which first protocol is NDP neighbor Discovery Protocol which is responsible for discovery of neighbors with in a network zone, it is initiated by the central node of zone on frequent basis to maintain updated list of all neighbor and possible routes. It broadcast hello beacon after specific time intervals to maintain status of connection to neighbors, it support the work of Intra Zone Routing Protocol which maintain connectivity with in one zone in network, which is protective in nature and follow link state routing technique. Other than IARP approach there is

one another protocol exist in ZRP technique which is used to perform communication outer-then a specific zone in network, that is known as Inter zone routing protocol IERP, it is reactive in nature and it initiates node discovery process outside zone, here one supporting protocol Border cast Resolution Protocol BRP helps to transmit message to all border nodes of current zone so that on receiving of this message that border nodes can transmit this message to its destination outer-then this zone with help of IERP.

VI. EVALUATION OF SIMULATION RESULTS

A. Simulation Environment

The Environment which is used for simulation[8] is a virtualized environment where one can simulate the desired network layout using virtual objects of all network elements in desired topology. It provides a platform to simulate network as well as find simulation results and evaluation tools to find out conclusions.

The simulation environment here used is famous network simulator which is known as NS2 [5][9] which is Linux operating system based simulation package It provide platform to simulate network, provide real visualize view of transmission process and also generates trace files. Ns-2 provides platform of routing of wired network and wireless networks. NS2 is consist of two type of working tools which are "ns" and "nam" also known as network simulator and network animator. It is useful to visualize network. It also has graph tools which can generate graphs after simulation process.

B. Traffic Model

CBR constant Bit Rate model have used for this work which is the source of traffic in simulated network [7][9]. Pairs of the source and destination nodes are distributed in a grid topology and it uses data packets of size 512-byte for simulation analysis. The packet rate of transmission can be varying time to time due to varying load of network.

C. Mobility Model

For the mobility model [5] the random way point of is used in this simulation. There is a configuration of 500 x 500 area plot for simulation in where all network nodes has configured in a grid topology. The Radio Model is two way ground and CBR as traffic and the packet size try to maintained as 512 bytes, network speed used is 10 m/s, and area for simulation is 500X500 wide area and number of nodes are 100 and simulation time are 20,40,60,80 & 100.

Simulation plan is to simulate network with OLSR and ZRP protocols and find out the simulation [10] performance variation between these two protocols in some scenarios.

VII. PERFORMANCE EVALUATION

Perform evaluation is based on the output of evaluation matrices which provide some parameters to specify the simulated network performance in various aspects like ratio of packet delivery, throughput of network, and analysis of packet transmission delay from one end to another end in

network[10]. After evaluation of these matrices some conclusion has found which shown in below graphs and tables.

A. Average Throughput

Network throughput [9][10] is ratio of messages delivered successfully from sender to receiver by a communication channel that can be physical or logical and it can pass by any network node. It has unit bits per second.

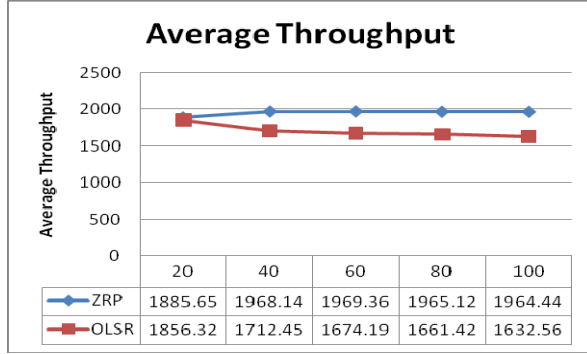


Fig. 1. Average Throughput

On the evaluation of simulation results, the value of network throughput of network using ZRP protocol in different simulation time initially higher as 1885.65 but it increases as simulation time increases as on 100 simulation time it reach on 1964.44 whereas, the throughput value of network based on OLSR based network in different simulation time initially 1856.32 but it decrease as simulation time increase as on 100 simulation time it reach on 1632.56 which shows that in ZRP based network can provide good network throughput.

B. Packet Delivery Ratio

It is the ratio of number of packets sent by sender and the number of packets received by receiver in transmission [9][10]. This ratio also measures the loss packets during transmission and also it is helpful in analysis of efficiency of routing protocol.

$$\frac{\sum \text{Number of packet receive}}{\sum \text{Number of packet send}}$$

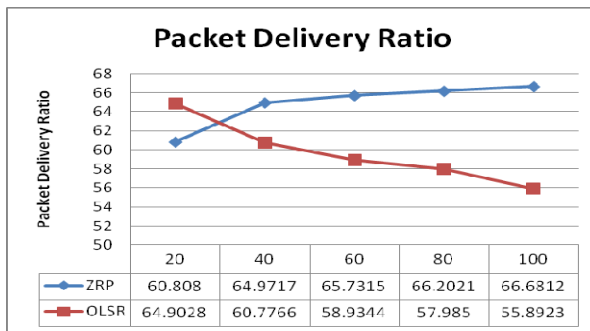


Fig. 2. Packet Delivery Ratio

On the evaluation of simulation results, packet delivery ratio evaluation of simulated network which is using ZRP

protocol in different simulation time initially low as 60.808 but it increases as simulation time increases as on 100 simulation time it reach on 66.6812 whereas, the packet delivery ratio of network based OLSR based network in different simulation time initially 64.9028, but it decrease as simulation time increase as on 100 simulation time and reach on 55.8923. Hence ZRP based network can provide good packet delivery ratio.

C. Average End-to-End delay

It is the time difference that is taken by transmitted packets to complete the travel start by sender and reach to receiver during transmission in network simulation [9][10]. This delay is a sum of all delay that is taken by the process of route discovery time, time taken for propagation.

$$\frac{\sum (\text{arrive time} - \text{send time})}{\sum \text{Number of connections}}$$

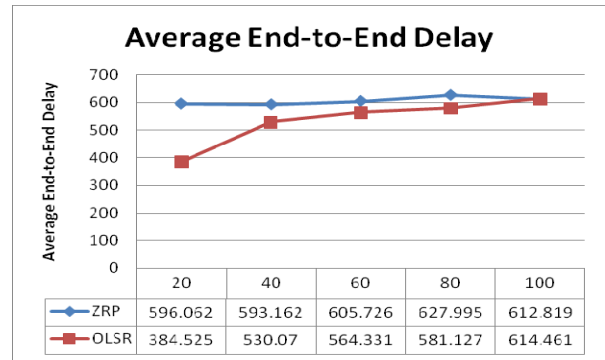


Fig. 3. Average End to End Delay

Based on simulation results, the average end to end delay in the network which is based on ZRP protocol in different simulation time initially low as 596.062 but it increases as simulation time increases as on 100 simulation time it reach on 612.819 whereas, the end to end delay of network based on OLSR protocol based network in different simulation time initially 384.525, but it increase as simulation time increase as on 100 simulation time and reach on 614.461. It shows that ZRP based network is capable to provide end to end delay.

VIII. CONCLUSION

Analysis of protocols in different network scenarios with varying data traffic shows that, OLSR and ZRP both work different in different scenarios and they are affected by varying data traffic due to varying simulation time in the network. It seems that ZRP is good performer instead of the OLSR protocol.

IX. FUTURE WORK

Future task could be enhancement of new features in ZRP protocol to make more effective protocol in the mobile ad hoc networks.

REFERENCES

- [1] Beigh Bilal Maqbool Prof.M.A.Peer "Classification of Current Routing Protocols for Ad Hoc Networks - A Review " International Journal of Computer Applications (0975 – 8887) Volume 7– No.8, October 2010
- [2] Asma Ahmed, A. Hanan, Shukor A. R., Izzeldin M. "Routing in Mobile Ad hoc Network " IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.8, August 2011
- [3] Sunil Taneja , Ashwani Kush "A Survey of Routing Protocols in Mobile Ad Hoc Networks" International Journal of Innovation, Management and Technology(2010 - 0248) Vol. 1, No. 3, August 2010
- [4] Md. Anisur Rahman, Md. Shohidul Islam, Alex Talevski "Performance Measurement of Various Routing Protocols in Ad-hoc Network " Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I, IMECS 2009, March 2009
- [5] Samyak Shah, Amit Khandre, Mahesh Shirole, Girish Bhole " Performance Evaluation of Ad Hoc Routing Protocols Using NS2 Simulation " Mobile and Pervasive Computing (CoMPC–2008)
- [6] Padmini Misra " Routing Protocols for Ad Hoc Mobile Wireless Networks" "http://www.cis.ohio-state.edu/~misra"
- [7] Furqan Haq and Thomas Kunz "Simulation vs. Emulation: Evaluating Mobile Ad Hoc Network Routing Protocols "Systems and Computer Engineering Carleton University Ottawa, Ont., Canada K1S 5B
- [8] Gianni A. Di Caro , " Analysis of simulation environments for mobile ad hoc networks " Technical Report No. IDSIA-24-03 IDSIA / USI-SUPSI , Dalle Molle Institute for Artificial Intelligence Galleria, Switzerland , December 2003
- [9] Karthik sadasivam " Tutorial for Simulation-based Performance Analysis of MANET Routing Protocols in ns-2 "
- [10] Kapang Lego, Pranav Kumar Singh, Dipankar Sutradhar "Comparative Study of Adhoc Routing Protocol AODV , DSR and DSDV in Mobile Adhoc Network" Indian Journal of Computer Science and Engineering Vol.1 No.4 364-371.
- [11] P. Jacquet , Le Chesnay, France ; P. Muhleth - Optimized link state routing protocol for ad hoc networks - Multi Topic Conference, 2001. IEEE INMIC 2001.
- [12] Nicklas, Beijar "Zone Routing Protocol (ZRP)" citeseer.nj.nec.com/538611.html.Wesley Longman Publishing Co., pp.221—253, (2001).

Current Moroccan Trends in Social Networks

Abdeljalil EL ABDOULI, Abdelmajid CHAFFAI, Larbi HASSOUNI, Houda ANOUN, Khalid RIFI
RITM Laboratory, CED Engineering Sciences
Ecole Supérieure de Technologie
Hassan II University of Casablanca, Morocco

Abstract— The rapid development of social networks during the past decade has led to the emergence of new forms of communication and new platforms like Twitter and Facebook. These are the two most popular social networks in Morocco. Therefore, analyzing these platforms can help in the interpretation of Moroccan society current trends. However, this will come with few challenges. First, Moroccans use multiple languages and dialects for their daily communication, such as Standard Arabic, Moroccan Arabic called “Darija”, Moroccan Amazigh dialect called “Tamazight”, French, and English. Second, Moroccans use reduced syntactic structures, and unorthodox lexical forms, with many abbreviations, URLs, #hashtags, spelling mistakes. In this paper, we propose a detection engine of Moroccan social trends, which can extract the data automatically, store it in a distributed system which is the Framework Hadoop using the HDFS storage model. Then we process this data, and analyze it by writing a distributed program with Pig UDF using Python language, based on Natural Language Processing (NLP) as linguistic technique, and by applying the Latent Dirichlet Allocation (LDA) for topic modeling. Finally, our results are visualized using pyLDAvis, WordCloud, and exploratory data analysis is done using hierarchical clustering and other analysis methods.

Keywords: distributed system; Framework Hadoop; Pig UDF; Natural Language Processing; Latent Dirichlet Allocation; topic modeling; pyLDAvis; wordcloud; exploratory data analysis; hierarchical clustering.

I. INTRODUCTION

Twitter and Facebook platforms that are part of people connected life are considered the most popular platforms. According to the latest statistics, there are 936 million daily active users just for Facebook, with 83% outside the USA [1], and 316 million monthly active users for Twitter, with 77% outside USA [2]. In some countries, these two platforms have grown very fast. For instance, in Morocco, the country concerned by our research work in this article, Facebook has grown by 590,000 Moroccan users between January and October 2011 [3]. While the number of Moroccan user accounts in Twitter reached 26,666 and their number of tweets reached 780,000 by month, thus occupying the third place in the Arab world [3]. If we talk about services that made these two platforms popular we can say that Twitter is user-friendly and allows users to write short messages that can be limited to 140 characters called “tweets” in which they can post links or share images. On the other hand, Facebook allows users to create personal profiles, add other users as friends, and exchange messages, including status updates, moreover, users can share photos, links, and personal thoughts.

These statistics encouraged us to lead a study that aims the analysis of messages published by Moroccan users, on these two platforms despite the difficulties quoted before. In this paper, we propose a detection engine of Moroccan social trends, that can handle the generated data by Moroccan users to create a text corpus useful to analyze and visualize the Moroccan society trends in a chosen period. To build our detection engine; we rely on the Hadoop framework, which is a distributed system, usually used to realize an infrastructure for storage and processing [4].

This infrastructure is composed of four parts. The first one is the data extraction part which handles the streaming of data, related to Morocco society, from both platforms Twitter and Facebook, and then stores the data in our distributed system using the HDFS storage model [5]. The second part handles the processing of collected data. It starts by converting tweets in JSON format using JAQL (JSON Query Language) [6], and then proceeds in search of pertinent information contained in these data. For Facebook, we use an API wrapper written in Java programming language to handle the JSON format and extract the posts and comments directly from the Graph API of Facebook. Then we apply a distributed program based on the Natural Language Processing [22] to this data by running a Pig UDF [7] written in Python language. The third part use the previous result to generate the LDA corpus [14]. The fourth part is composed of visualization tools, such as pyLDAvis, WordCloud, and other exploratory data analysis like hierarchical clustering.

This paper is organized as follows. In Section II, we introduce some related works. In Section III, we present the tools and methods used in our system. In Section IV, we describe the architecture of the detection engine of Moroccan social trends. In section V, we run an experiment. We end with a conclusion in Section VI.

II. RELATED WORK

The analysis of social network platforms has been the focus of interest of many researches. For example, H. Kwak, C. Lee, H. Park and S. Moon [8] found that Twitter users sometimes broadcast news before traditional media. In 2011, J. Weng and B.-S. Lee [9] proposed a method based on the frequency of terms presented daily in the corpus. The frequency of each term is represented as a signal. More so, O. Ozdikiş, P. Senkul and H. Oguztuzun [10] performed a semantic expansion of the terms presented in the tweets. Finally, H. Becker, M. Naaman and L. Gravano [11], proposed an approach to identify, among

all the tweets, those describing events. They grouped all tweets according to their textual similarity.

On the other hand, many other research works focused on the detection and interpretation of emotion in Twitter platform. For example, V. Nguyen, B. Varghese and A. Barker [12], proposed a Framework to analyze and visualize public sentiment in the Twitter platform. The research presented in this article is based on a dictionary of words and a learning algorithm that detects feelings of the public in a specific region. In the same field, Hand, W. Wang, L. Chen, K. Thirunarayan and A. Sheth [13], performed self-extracting hashtags related to emotions that exist in Twitter (2.5 million tweets), and they have applied two learning algorithms to identify public emotions.

III. TOOLS AND METHODS

A. The Hadoop Framework

Our detection engine of Moroccan social trends is built by using a special infrastructure, based on the Hadoop Framework. The Apache Hadoop is an open-source software framework written in Java for distributed storage and distributed processing of very large data sets on computer clusters built from commodity hardware [4]. The Hadoop Framework consists of two primary components. The first one is HDFS [5], which stands for Hadoop distributed file system; it is an open-source data storage, inspired by GFS (Google File System) and stores large files across multiple machines. The second one is MapReduce, which is an open-source programming model developed by Google Inc. Apache adopted the ideas of Google MapReduce and developed it. MapReduce allows the decomposition of tasks and the integration of results.

The HDFS system can store files whose size can reach the terabytes. The stored files are then divided into blocks of 64 MB by default; these blocks are then replicated in three copies by default across the different nodes in the cluster. The HDFS uses a master-slave architecture, and it consists of:

a) Single NameNode, which is running on the master node and holds the metadata of HDFS by mapping data blocks to data nodes, it manages file system namespace operations like opening, closing, and renaming files and directories. Furthermore, it regulates access to these files by clients.

b) Secondary NameNode allows creating the checkpoints of the file system present in the Namenode.

c) DataNodes run on slave nodes, they are responsible for storing blocks within the node. They report every file and block stored in that node to the NameNode, and do all file system operations like creating, deleting files according to instructions received from the NameNode.

The key traits of HDFS are the scalability and availability due to data replication and fault tolerance.

The MapReduce is a program model for distributed computing based on Java modeled after Google's paper on

MapReduce [21]. It allows to parallelize processing over a large amount of stored data by decomposing data submitted by a client into small parallelized map and reduce workers. The map takes a set of data as a key-value pair and converts it into another set of data as a key-value pair. Then the reduce task takes the output from a map task as an input and combines those data tuples into a smaller set of tuples. The main work of MapReduce can be seen in between the reduce input and the map output, and the shuffle and sort stage. The MapReduce uses a master-slave architecture, and it consists of:

a) JobTracker, running on the master node and receiving a request from a client then assigning tasks to workers running on slave nodes where the data is locally stored, these tasks are performed by the TaskTrackers. If the TaskTracker, for some reason, fails to execute the job, the JobTracker assigns the task to another TaskTracker where the data are replicated.

b) TaskTracker is a daemon that is running on slave nodes, it starts and monitors the Map and Reduce tasks assigned by the JobTracker. The TaskTrackers report their status to the JobTracker by a heartbeat, and they send the free slots within it to process data.

MapReduce provides the following features; simplicity of development, scalability, automatic parallelization and distribution of work, fault tolerance.

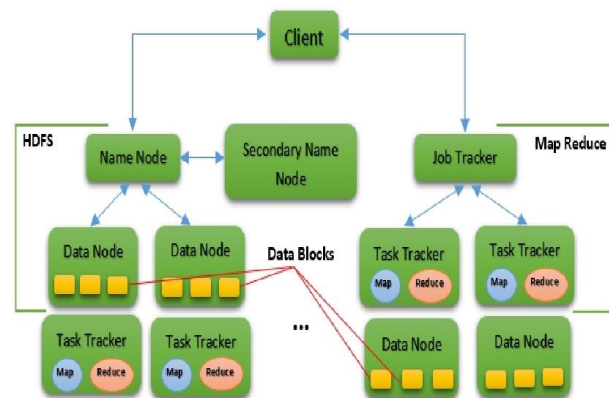


Figure 1. Hadoop Architecture

The use of Hadoop framework is fundamental in our system due to the complexity of used algorithms. It allows us to accelerate the processing and get better performance with high reliability.

B. The Natural Language Processing (NLP)

The Natural Language Processing [22] is the application of knowledge of languages in the development of intelligent computer systems that can recognize, understand, interpret and reproduce human language in its different forms. The new approaches of NLP are based on machine learning by using the concepts and techniques of artificial intelligence (AI) that

allows computers to learn without being programmed. The input and the output of an NLP system can be composed of speech or written texts.

The NLP system involves two main components. The first one is the Natural Language Understanding (NLU) that handles the machine reading comprehension and try to understand the meaning of a written text. The other component is the Natural Language Generation (NLG) that includes computational linguistics and allows computers to writes text of the same quality as that of a human being.

The NLP system is based on five phases [19]:

- a) **Lexical Analysis:** is concerned with the structure of words. To analyze this structure, the lexical analysis divides a text into sentences, and words.
- b) **Syntactic Analysis:** is concentrated on the analysis of the sentence structure by analyzing the words and their relationship within this sentence.
- c) **Semantic Analysis:** is interested in the meaning of sentences considered individually.
- d) **Discourse Integration:** handles the relationship between the current sentence and the sentence just before it, to give meaning to the current one.
- e) **Pragmatic Analysis:** determines the meaning of the text in context.

Using NLP we realize the following operations:

- Segmentation: Cut the text into "segments" (tokens).
- Part-of-speech tagging: associate to each word its appropriate syntactic category (noun, verb, adjective, etc.)
- Named entity extraction: classify words in a text into predefined categories; like company names, person or other things.
- Normalization: e.g. eliminate capital letters.
- Make statistical calculations on words.
- etc.

In our system, The NLP is used to process the raw data collected from the social networks Twitter and Facebook after the extraction of pertinent information using JAQL [6]. The processing contains successive steps that allow analyzing the Moroccan user-generated data expressed with different languages and dialects, used by Moroccan people when communicating on these social networks.

C. Topic models: LDA

Topic models are a set of algorithms that uncover and generate the hidden topics within a set of documents based on the word frequency. Many types and implementations of topic modeling have been published. The most common and the most used in search engines is Latent Dirichlet Allocation (LDA) [14]. LDA can also be used to explain the intent embedded in text or resolve ambiguous words in systems like search engines.

Topic models based on LDA work as a statistical machine learning and text data mining, which consist of a Bayesian inference model that calculate the probability distribution over topics in each document, where each topic is characterized by a probabilistic distribution based on a set of words (or n-grams, or syntactic n-grams). The LDA model needs the number of topics as a parameter to choose the number of partitions to divide the corpus.

For its computation, the LDA model uses as input tf-idf (term frequency-inverse document frequency) [20] values from documents by calculating the importance of each word, and the number of topics given by a user. After this, the generative model determines his probability of the membership of a document for each topic, generating new Vector Space Model of LDA topics [17], as shown in Figure 2.

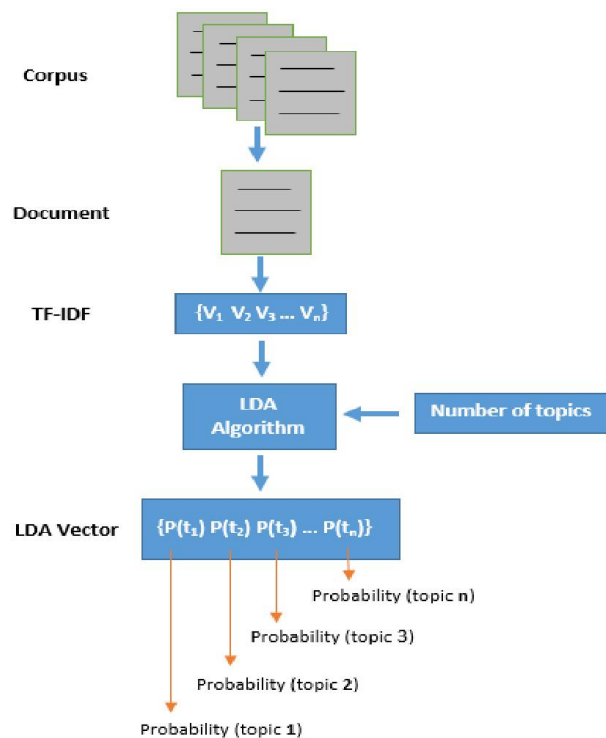


Figure 2. LDA Model

The LDA algorithm contains the following steps:

- 1) For each document:
 - Decide on the number of words N the document will have using the Poisson distribution, i.e.; only the words that correspond to this distribution are considered.
 - Choose a topic mixture for the document by using the Dirichlet distribution in order to ascertain the possibility of the membership of the document on each topic.
- 2) Generate each word in the document by:
 - Picking up a topic by using an iterative process; for each word w_i in the document, the algorithm chooses the topic t_n using multinomial distribution and conditional multinomial probability $P(w_i/t_n)$ [17].
 - Finally, using the topic to generate the word itself. The algorithm returns a new set of words for each topic and the probability of each document to belong to each topic.

We decide to use the LDA model in our system to detect the preoccupations and interests related to the Moroccan society, and to discover new social trends related to it. For this reason, we implement a free python library for topic modeling based on LDA model called "gensim".

D. Principal component analysis (PCA)

Using topic model like LDA allows detecting and interpreting the hidden topics in a corpus. However, when the number of topics increase and becomes high, we can find difficulties in distinguishing between topics and interpreting the results correctly. Reducing the number of topics is sometimes not an option, but becomes a necessity. To solve this issue, we use the Principal Component Analysis (PCA) [18].

PCA is a powerful tool for analyzing and expressing data in such a way to detect the similarities and differences. It involves a mathematical procedure that helps revealing interrelationships among a large number of variables. PCA can transform some correlated variables into a smaller number of uncorrelated variables called *principal components*, which are a new variable, obtained as linear combinations of the original variables. The first principal component is required to have the largest possible variance. The second component is measured based on the fact that it is orthogonal to its first component and having the largest Inertia, which is the sum of the squared element of a variable:

$$v_j^2 = \sum_i x_{ij}^2$$

The other components are computed likewise. The values of these new variables for the observations are called factor scores, which can be interpreted geometrically as the projections of the observations onto the principal components.

Using the PCA in our system, we managed to compress the size of data by reducing the number of dimensions through the use of the probabilities of each word in each topic. Our goal is to keep the important information and to make the visualization of data clear and understandable without much loss of information.

E. Hierarchical Clustering: Dendrogram

The hierarchical clustering [23] is a method of group analysis that allows creating a hierarchy of clusters. In other words, the main purpose of this method is to group objects of similar kind into clusters, where objects of each cluster are close to each other. This is done by repeating the calculation of distance measures between objects, and between clusters to construct a tree, and then the result can be represented graphically as a dendrogram.

Taking N as the number of objects, the hierarchical clustering needs a set of object-to-object distances defined as $N * (N - 1) / 2$ and a cluster-to-cluster distances computed with linkage function [24]. The algorithm of the hierarchical clustering can be outlined as follows:

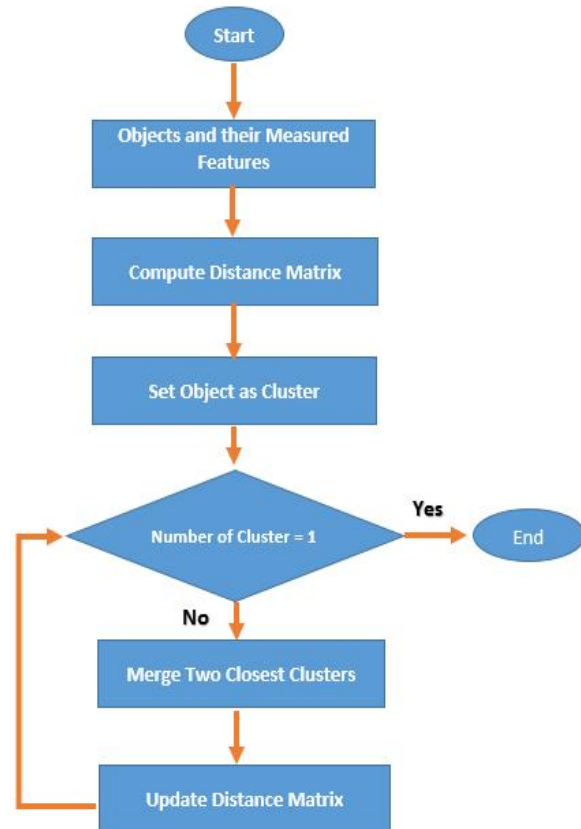


Figure 3. Hierarchical Clustering Algorithm

There are two main conceptual approaches for the Hierarchical Clustering: Hierarchical agglomerative clustering and Divisive clustering. The first one is a bottom-up clustering method, it starts with every single object in a single cluster, and then in each successive iteration, it merges the similar clusters until all objects belong to one cluster. The second one is a top-down clustering method; it works in a similar way to agglomerative clustering but in the opposite direction, it starts with one cluster containing all objects and then in each successive iteration, splits resulting clusters into smaller clusters until each cluster contain a single object. In order to combine similar clusters (for agglomerative), or split one cluster into similar clusters (for divisive), a measure of dissimilarity between sets of observations is required, this is obtained by calculating the distance between pairs of observations.

We use the hierarchical clustering analysis in our system to transform the *topic * word* matrix into topic-by-topic distance matrix, in order to group the correlated topics together. The result of hierarchical clustering is similar to PCA, except we can observe the hierarchical ordering of topics.

F. Force-directed graph

The Force-directed graph is a set of algorithms widely used for drawing efficient graphs due to their flexibility, ease of implementation, and the aesthetically pleasant drawings they produce. These graphs are generally visualized as node-link diagrams, in which dots depict the nodes, joined by lines for the edge. The Force-directed graph is used in several use cases such as traffic networks, social networks, software systems, etc.

Using the same distance matrix computed in the hierarchical clustering, we can plot the topics using the Force-directed graph. The nodes represent topics and edges represent the topic covariance metric, in consequence, we can observe the relationship between topics according to the distance between them.

In our system, NetworkX [15] handles the visualization of topics using the Force-directed graph, which is a Python package for the creation and handling of complex networks. We choose the spring layout of Force-directed graph that pulls linked and similar nodes toward the center.

IV. ARCHITECTURE OF THE DETECTION ENGINE OF MOROCCAN SOCIAL TRENDS

Our Detection Engine of Moroccan Social Trends (DEMST) is composed of four parts as follow:

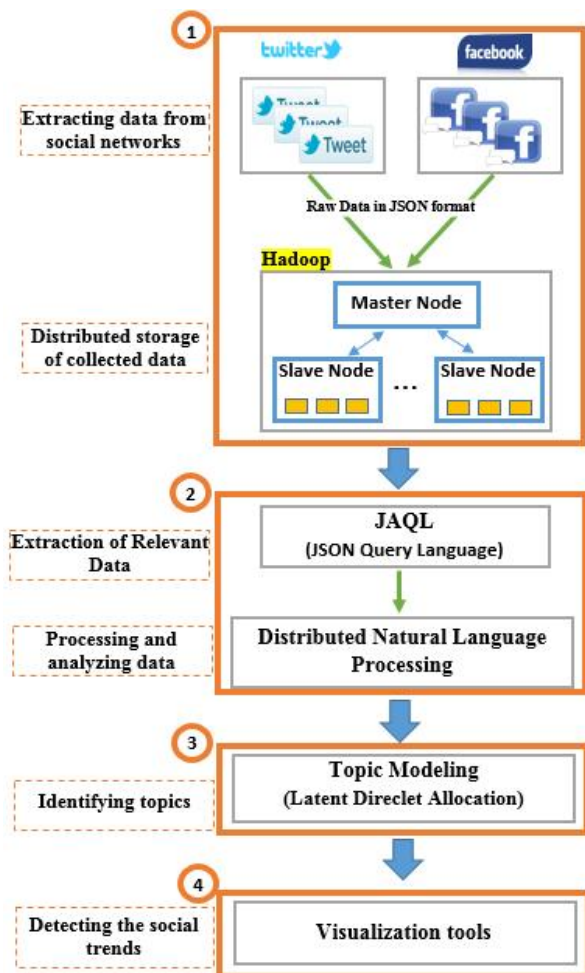


Figure 4. Architecture of DEMST

A. Extraction and distributed storage of raw data

The first part of DEMST involves the extraction of data from the social networks Twitter and Facebook because this is the most important task in the data analysis process. All these data are related to the Moroccan society and reflect the concerns of Moroccan users. This data is stored in our system using the Hadoop distributed file system (HDFS).

1) Streaming Data from Twitter

Comparing to the other social network platforms, Twitter API is more open and accessible. Three different ways are used to access Twitter data and return well-structured tweets in JSON format that facilitate analysis and access to the desired information. Returned tweets contain a variety of information: text, hashtags, the number of times retweeted, the information of the user, etc.

The three ways offered by Twitter to get tweets are Search API, Streaming API, and Twitter firehose:

a) **Search API**: allows to query data on Twitter using some criteria such as the keywords, locations, usernames, etc.

b) **Streaming API**: provides tweets that happen in near real-time by launching a request for an unlimited period as long as no problem occurs.

c) **Twitter firehose**: similar to the Streaming API and provides methods of writing and reading, which allows access to the Twitter database. This includes the extraction of the latest tweets sent on Twitter

To access Twitter data we need to be authorized. After registering on Twitter and creating an account on <https://apps.twitter.com>, Twitter provides us four secret keys: consumer key, consumer secret key, access token and access token secret. Once we have the keys, we can use the Streaming API in our system to retrieve tweets.

For the extraction of related tweets to Moroccan society, we used the geolocation available in the new version 1.1 of Twitter API that allows us to filter tweets by location. For this reason, we used the <http://boundingbox.klokantech.com> website to determine the geographical coordinates (latitude and longitude) of Morocco. Our system handles the streaming of data from Twitter using a library of Python language called Tweepy [25], which allows access to Twitter API.

```
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
stream = tweepy.Stream(auth, l)
stream.filter(locations=[-17.2122302, 21.3365321, -0.9984289, 36.0027875], async='true', encoding='utf8')
```

To ignore the tweets from neighboring countries of Morocco such as Algeria, Mauritania, Spain and Portugal, we have had recourse to the code of Morocco which is 'MA':

```
if decoded['place']['country_code'] != 'MA':
```

2) Streaming data from Facebook

The Facebook platform provides a set of services and tools allowing developers to create applications that access data in Facebook. This data is rich and contain sensitive information about users; that is why Facebook apply restrictions to the private accounts and make the access controls complicated.

Using this platform, the developers can integrate massive data of Facebook called social graph and get insights on the social interactions, using the Graph API.

The Graph API of Facebook is a low-level HTTP-based API that offers to developers the ability to query data, manage photos, post a new status message and other tasks. The Graph API provides a search functionality similar to Twitter. For example, it allows finding places, events, public posts, users, etc. Like Twitter, the graph API return well-structured data in JSON format.

The use of the Graph API of Facebook involves the authentication with Facebook App ID, Facebook App Secret, and Facebook OAuth token. Using the URL <https://developers.facebook.com/apps> we can obtain the App ID and App Secret. And the OAuth token can be obtained by the Graph API Explorer using the URL <https://developers.facebook.com/tools/explorer/>. The OAuth token is generated for every user. It identifies both the user and the application, and manages the permissions.

Using *Facebook4j* [26], a Facebook API wrapper written in Java, we can access to the Graph API of Facebook using the generated token.

```
import facebook4j.*;
Facebook fbk = new FacebookFactory().getInstance();
fbk.setOAuthAppId(FBAppID, FBAppSecret)
fbk.setOAuthAccessToken(new AccessToken(FBAccessTokens));
```

Due to the restrictions applied by the Facebook Platform On private accounts, our system is limited to use the data of the public pages maintained by the Moroccan users. For this purpose, we utilize *SocialBakers*, a popular provider of social media analytic tools, statistics, and metrics to get the IDs of the most popular Facebook pages about Morocco which have the best ranking and a huge number of subscribers. This method allows us to retrieve the posts and comments of Moroccan users on every page obtained by the ID.

```
ResponseList<Post> feeds =
fbk.getFeed(listPopularPagesIDMorocco[]);
for (int i = 0; i < feeds.size(); i++) {
    Post post = feeds.get(i);
    if (post.getMessage() != null) {
        outputStream.writeUTF(post.getMessage());
    }
    PagableList<Comment> comments =
post.getComments();
Iterator<Comment> it = comments.iterator();
while (it.hasNext()) {
    outputStream.writeUTF(it.next().getMessage());
}
}
```

3) Distributed storage using Hadoop

The collected data from Twitter and Facebook are stored directly in our distributed infrastructure using the HDFS of Hadoop system.

The storage of tweets, retrieved from the Twitter platform, in our distributed system is handled by *Hadoopy* [27], which is a Python wrapper for Hadoop using Cython. Hadoopy allows to read and write data directly to HDFS from Python. To write our program, we use The IPython Notebook, which is an interactive environment for python.

```
import hadoopy
```

```
hdfs_path = 'hdfs://master:54310/tweets/'
```

```
hadoopy.put(Data ,hdfs_path)
```

Using Eclipse, an IDE for Java and after adding the necessary JAR files of Hadoop, we wrote a program that handles the storage of posts and comments retrieved from the Facebook platform.

```
import org.apache.hadoop.conf.*
```

```
import org.apache.hadoop.fs.*
```

```
Configuration conf = new Configuration();  
conf.addResource(new Path("/home/hduser/hadoop/conf/core-site.xml"));
```

```
conf.addResource(new Path("/home/hduser/hadoop/conf/hdfs-site.xml"));
```

```
FileSystem fs = FileSystem.get(conf);  
FSDataOutputStream outputStream = fs.create(file);  
outputStream.writeUTF(post.getMessage());
```

B. Extraction of pertinent information with JAQL and processing data with NLTK

The second part concerns the treatment of the collected data, starting with parsing raw data and retrieving relevant information using Jaql (JSON Query Language), and then applying a distributed program based on natural language processing to analyze the parsed data

1) Parsing raw data with JAQL

JAQL is a functional data processing and a query language designed for JavaScript Object Notation (JSON), a highly used data format known of its simplicity and modeling flexibility. It was donated by IBM to the open source community. JAQL is primarily used to process large-scale structured and deeply nested semi-structured data, and it can be run to query big data inside HDFS using Framework Hadoop.

As we said earlier, the APIs of Twitter and the Graph API of Facebook utilize the JSON format to respond to queries. For this reason, we use JAQL to parse and extract the relevant data from the tweets. The fields obtained are:

- **Created_at**: the creation date of the tweet.
- **Lang**: language used to write the tweet (necessary for the processing with NLP).
- **Text**: the body of the tweet that contains the personal thoughts of the user.

To write the program of parsing raw data, we have utilized the Eclipse IDE for Java, after having added the necessary JARs of both Hadoop and JAQL.

```
JaqlQuery jaql = new JaqlQuery();
```

```
jaql.setQueryString("read(hdfs($location),{format:'org.apache.hadoop.mapred.TextInputFormat', converter:'com.ibm.jaql.io.hadoop.converter.FromJsonTextConverter'}) -> transform {$.created_at,$.text,$.lang}");
```

```
JSONArray jv = (JSONArray) jaql.evaluate();
```

```
JsonParserFactory factory=JsonParserFactory.getInstance();
```

```
JSONParser parser=factory.newJsonParser();
```

```
Map jsonData=parser.parse.Json(jv.get(0).toString());
```

```
String at=(String)jsonData.get("created_at");
```

```
String text=(String)jsonData.get("text");
```

```
String lang=(String)jsonData.get("lang");
```

Facebook4j, the wrapper for The Facebook Graph API, allows us to avoid parsing Facebook's JSON responses with JAQL because it already encapsulates the desired result in a Java object, so we simply need access to the object that have the desired data.

```
Post post = feeds.get(i);
```

```
post.getComments();
```

2) Processing data with NLP

The core of our detection engine of Moroccan social trends is the NLP; it contains all the necessary algorithms to handle the linguistic diversity of our Moroccan society.

We choose to work with the Natural language processing Toolkit (NLTK), which is a suite of open-source Python modules, allowing programs to work with the human language data. It provides over 50 corpora and lexical resources such as WordNet and a set of text processing libraries for tokenization, parsing, classification, stemming, etc.

The steps we followed to treat the collected data from Twitter and Facebook are:

a) Delete unnecessary data: usernames, emails, hyperlinks, retweets, punctuation, possessives from a noun, duplicate characters, and special characters like smileys.

b) Normalize whitespace (strip whitespace from the start and end of a word and convert multiple sequential whitespace chars into one whitespace character).

c) Convert hashtags into separate words, for instance, the hashtag #MoroccanSociety is converted into two words Moroccan and society.

d) Transform words of Moroccan dialect, or in a dialect of Berber Tamazight into the standard Arabic. These words could be written using the French or Arabic alphabet. to perform this task, we create a dictionary of words that we gathered in a python file which we store in each slave node of our cluster, and imported in the NLP script executed in these nodes. Below, an example of this file

```
#-*- coding: utf8 -*-
moroccanDialect = [
('katbghi', u'تخب'),
('khas', u'يجب'),
('ban', u'يظهر'),
...
(u'إبيه', u'نعم'),
(u'ارجوك', u'عافاك'),
(u'امسك', u'خود'),
...
('zgizzi', u'يتجادل'),
('zigiz', u'يشغل'),
('werg', u'يحمل')
```

e) Create a function to detect automatically the language used in the text (Standard Arab, French or English).

f) Create a function for automatic correction of words written in standard Arabic, French and English.

g) Create a list of contractions to normalize and expand words like What's=What is

h) Stemming: Delete the suffix of a word written in standard Arabic, French or English until we find the root.

i) Identify and remove tokens of part of speech that are irrelevant to our analyze using a software of Stanford University called Part-Of-Speech Tagger (POS Tagger). This software reads the text in Standard Arabic, French or English, and assigns parts of speech to each word such as noun, verb, adjective, etc.

j) Remove stopwords for standard Arabic (بعد, إن, أن, ...), French (alors, à, ainsi, ...), and English (about, above, almost, ...).

These steps are assembled in two python files; one for the tweets and the other for posts and comments of the Moroccan pages on Facebook called successively NLTK_Tweet.py and NLTK_FB.py. These two python files are processed in a distributed manner by the Apache Pig, which is a high-level scripting language for processing and analyzing large data sets using Hadoop MapReduce.

The Apache Pig allows to create user-defined functions using programming languages like Java and Python to specify a custom processing. The python files NLTK_Tweet.py and NLTK_FB.py are processed successively by the Pig files Pig_Tweet.pig and Pig_FB.pig. The script of the python files need to be registered in the script of the Pig files using Streaming_python as follow:

```
REGISTER 'hdfs://master:54310/apps/NLTK_Tweet.py' USING
streaming_python AS nltk_udfs;
```

Then we load data and call our function defined in the file of python script by using:

```
data = LOAD '/socialData/*' using TextLoader() AS
(line:chararray);
```

```
Result = FOREACH data GENERATE
nltk_udfs.NameFunction(line);
```

3) The Apache Oozie Workflow

The Apache Oozie is a workflow scheduler system to manage Apache Hadoop jobs. A workflow is a series of actions forming a Directed Acyclical Graphs.

Using Oozie Workflow, we can chain the parsing of data by JAQL with the processing of data with the NLP. This workflow allows us to choose a particular period to execute our system and simplifies its use.

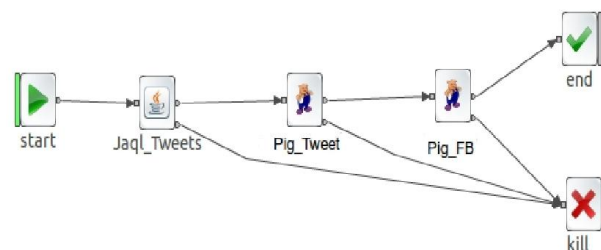


Figure 5. Oozie Workflow

C. Generating the LDA corpus

The third part of our system uses the previous result of the workflow composed of JAQL and NLP to produce the LDA corpus. LDA is a probabilistic model used to determine the covered topics using the word frequency in the text.

To generate the LDA model, we need to construct a document-term matrix with a package called "Gensim", which allows us to determine the number of occurrences of each word in each document.

```
from gensim import corpora, models

texts = [word.split() for word in documents.split(",")]

dictionary = corpora.Dictionary(texts)
```

The *Dictionary()* function take the text as input and assign a unique integer id to each unique word while also collecting word counts and relevant statistics. This dictionary must be converted into a bag-of-words using:

```
corpus = [dictionary.doc2bow(text) for text in texts]
```

The *doc2bow()* function converts the dictionary into a bag-of-words in the form of a list of tuples (word_id, word_frequency) for each document. The result, corpus, is a list of vectors equal to the number of documents, where each vector is a series of tuples.

After creating the document-term matrix, we can generate the LDA model using the *LdaModel* class:

```
lda=models.Ldamodel.LdaModel(corpus,id2word=dictionary,
num_topics=x , passes=x)
```

The parameters utilized in this class are:

- **id2word:** Our previous dictionary. Required by The LDA Model class to map ids to string.
- **num_topic:** The number of topics to generate by the LDA model.
- **passes:** This value is optional and is used by the model to know the number of laps through the corpus. This number of passes determines the accurate of the model.

To show topics detected by our LDA model we use:

```
lda.show_topics()
```

The result has the format below:

```
0.014*word1 + 0.009* word2 + 0.009* word3 + 0.006* word4 +
0.016*word5 + ...
```

D. Visualizing the topics

The last part involves getting insights from our LDA model by using methods and tools allowing the interpretation and visualization of discovered subjects. The first visualization tool we used is the WordCloud, which gives us an idea about dominant topics in our corpus. We install the WordCloud package using the command *pip install wordcloud*. This package utilizes the frequencies of words calculated by the LDA model to graphically display these words as shown in the following script:

```
import wordcloud

scores = [float(x.split("**")[0]) for x in final_topics.split(" + ")]

words = [x.split("**")[1] for x in topics.split(" + ")]

freqs = []

for word, score in zip(words, scores):

    freqs.append((word, score))

elements = wordcloud.fit_words(freqs, width=1800, height=1000)

wordcloud.draw(elements, "WordCloud.png",width=1800,
height=1000, font_path="Thabit.ttf")
```

The second tool is *pyLDAvis*, an interactive web-based visualization tool, which retrieves information from the LDA topic model and represents the results in an attractive way to help users to interpret the detected topics. The *pyLDAvis* can be installed using the command *pip install pyldavis*; this package is intended to be used in the IPython notebook. We can visualize the result by running the script below:

```
import pyLDAvis.gensim as gensimvis

import pyLDAvis

vis_data = gensimvis.prepare(lda, corpus, dictionary)

pyLDAvis.display(vis_data)
```

To reduce the number of topics and increase their precision, we use the PCA technique for the visualization, which allows us to project data into two dimensions and merge correlated topics in a single theme.

Before applying the PCA technique on the topics, we need to transform the categorical data into numerics by affecting indicator variables for each category. We use the scikit-learn, an open-source and efficient tool for data mining and data analysis, to accomplish the data transformation and the PCA technique. The next script uses the function *DictVectorizer()* to directly convert strings into numeric values in order to be used in the PCA technique as following:

```
from sklearn.feature_extraction import DictVectorizer
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt

def topics_to_vectorspace(i, n_words=100):
    rows = []
    for i in xrange(n_topics):
        temp = lda.show_topic(i, n_words)
        row = dict(((i|1],i|0]) for i in temp)
        rows.append(row)
    return rows

vec = DictVectorizer()
X = vec.fit_transform(topics_to_vectorspace(n_topics))
X.shape

pca = PCA(n_components=2)
X_pca = pca.fit(X.toarray()).transform(X.toarray())
plt.figure()
for i in xrange(X_pca.shape[0]):
    plt.scatter(X_pca[i, 0], X_pca[i, 1], alpha=.5)
    plt.text(X_pca[i, 0], X_pca[i, 1], s=' ' + str(i))

plt.title('PCA Topics')
plt.savefig("pca_topics")
plt.close()
```

Another way to visualize topics is the hierarchical clustering using the dendrogram, a tree-structured graph that represents the relationships of similarity among the topics. This similarity is calculated using the Euclidean distance:

```
from scipy.spatial.distance import pdist, squareform
import matplotlib.pyplot as plt

plt.figure()
corr = squareform(pdist(X.toarray(), metric="euclidean"))
plt.figure(figsize=(12,6))
R = dendrogram(linkage(corr))
plt.savefig("similarity_topics")
plt.close()
```

The last visualization tool we used in our is the Force-directed graph using the package NetworkX of Python language. The nodes represent the topics, and the edges represent the relationship between them, so the similar topics are closer to each other. We use the Euclidean distance to calculate this similarity:

```
import networkx as nx

corr = squareform(pdist(X.toarray(), metric="euclidean"))

G = nx.Graph()
for i in xrange(corr.shape[0]):
    for j in xrange(corr.shape[1]):
        if i == j:
            G.add_edge(i, j, {"weight":0})
        else:
            G.add_edge(i, j, {"weight":1.0/corr[i,j]})

edges = [(i, j) for i, j, w in G.edges(data=True) if w['weight'] > .8]

edge_weight=dict(((u,v),int(d['weight'])) for u,v,d in
G.edges(data=True))

pos = nx.spring_layout(G)
nx.draw_networkx_nodes(G, pos, node_size=100, alpha=.5)
nx.draw_networkx_edges(G, pos, edgelist=edges, width=1)
nx.draw_networkx_labels(G, pos, font_size=8, font_family='sans-serif')

plt.savefig("network")

plt.close()
```

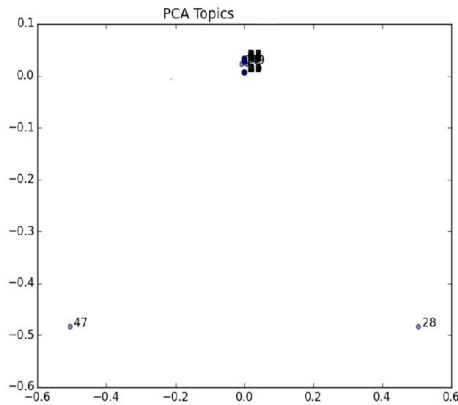
V. EXPERIMENTS AND RESULTS

We experimentally evaluated our detection engine of Moroccan social trends by running it and see if it can help us to understand what happening in our society by detecting the current trends.

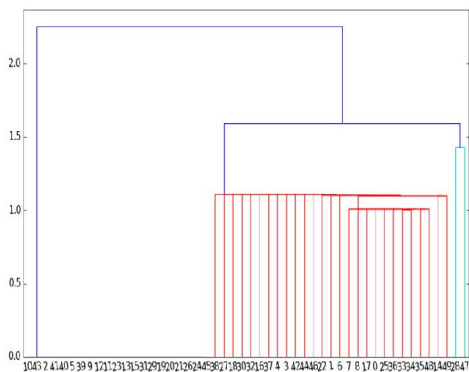
A. Data

During the streaming of tweets from the Platform Twitter, we examine the content of the tweets to verify if the source is Morocco, and we found that all the tweets without exception have the same code of country which is "MA" indication for Morocco country.

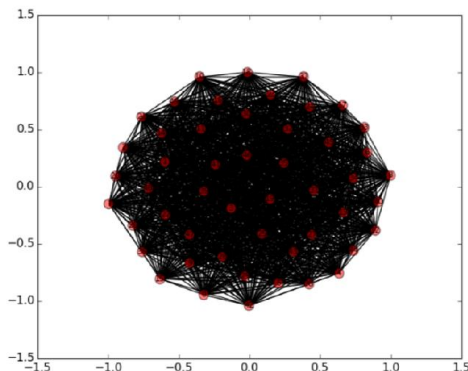
3) PCA



4) Hierarchical clustering



5) Force-directed graph



This experiment reveals that Moroccan users show a high interaction with the earthquake subject that recently hits the north part of the country. This was expressed by the usage of

expressions (by users) like earthquake scale, “Hosima”, “Nador”, which are the location of the incident.

VI. CONCLUSION AND FUTURE WORK

The detection engine of Moroccan social trends developed in this paper is a reliable system that allows to collect, analyze, and visualize Moroccan social trends in Twitter and Facebook platforms. The growing number of users was an influential factor that creates a large sample size that can be used to detect the current trends. It has to be noted that the system still needs some improvement because the growing usage of different languages. Also, the Moroccan dialect is liquid which can lead to addition or change of meaning of some words. Therefore, enhancing the analyzes of Moroccan dialect will require more future improvement.

VII. REFERENCES

- [1] Socialbakers.com, "All Facebook statistics in one place". [Online]. Available: <http://www.socialbakers.com/statistics/facebook>. [Accessed: 01- Jan- 2016].
- [2] Socialbakers.com, "All Twitter statistics in one place". [Online]. Available: <http://www.socialbakers.com/statistics/twitter>. [Accessed: 01- Jan- 2016].
- [3] Lemag : The Maghreb Daily, "Facebook and Twitter in Morocco". [Online]. Available: http://www.lemag.ma/english/Facebook-and-Twitter-in-Morocco-key-figures_a952.html. [Accessed: 01- Jan- 2016].
- [4] Wikipedia, "Apache Hadoop". [Online]. Available: https://en.wikipedia.org/wiki/Apache_Hadoop. [Accessed: 04- Jan- 2016].
- [5] Mrudula Varade and Vimla Jethani, "Distributed Metadata Management Scheme in HDFS", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.
- [6] Code.google.com, "Google Code Archive - Long-term storage for Google Code Project Hosting.". [Online]. Available: <https://code.google.com/p/jaql/wiki/JaqlOverview>. [Accessed: 03- Jan- 2016].
- [7] Pig.apache.org, "User Defined Functions". [Online]. Available: <https://pig.apache.org/docs/r0.9.1/udf.html>. [Accessed: 06- Jan- 2016].
- [8] H. Kwak, C. Lee, H. Park and S. Moon, "What is Twitter, a social network or a news media?", *Proceedings of the 19th international conference on World wide web - WWW '10*, 2010.
- [9] J. Weng and B.-S. Lee, "Event detection in twitter", *In 5th In. AAI Conf. on Weblogs and Social Media*, 2011.
- [10] O. Ozdikian, P. Senkul and H. Oguztuzun, "Semantic Expansion of Tweet Contents for Enhanced Event Detection in Twitter", *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2012.
- [11] H. Becker, M. Naaman and L. Gravano, "Beyond trending topics: Real-world event identification on twitter", *In Proceedings of the Fifth International AAI Conference on Weblogs and Social Media (ICWSM11)*, 2011.
- [12] V. Nguyen, B. Varghese and A. Barker, "The royal birth of 2013: Analysing and visualising public sentiment in the UK using Twitter", *2013 IEEE International Conference on Big Data*, 2013.
- [13] W. Wang, L. Chen, K. Thirunarayan and A. Sheth, "Harnessing Twitter "Big Data" for Automatic Emotion Identification", *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, 2012.
- [14] D. Blei, A. Ng, and M. Jordan, "Latent Dirichlet allocation", *Journal of Machine Learning Research*, 2003.

- [15] Networkx.github.io, "Overview — NetworkX", 2016. [Online]. Available: <https://networkx.github.io/>. [Accessed: 09- Jan- 2016].
- [16] R. Vijayakumari, R. Kirankumar and K. Gangadhar a Rao, "Comparative analysis of Google File System and Hadoop Distributed File System", *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 3, 2014.
- [17] V. Carrera - Trejo, G. Sidorov, S. Miranda - Jiménez, M. Moreno Ibarra and R. Cadena Martínez, "Latent Dirichlet Allocation complement in the vector space model for Multi - Label Text Classification", *International Journal of Combinatorial Optimization Problems and Informatics*, vol. 6, 2015.
- [18] H. Abdi and L. J. Williams, "Principal component analysis", 2010
- [19] A. Chopra, A. Prashar and C. Sain, "Natural Language Processing", *International Journal of Technology Enhancements and Emerging Engineering Research*, vol. 1, 2013.
- [20] Z. Yun-tao, G. Ling and W. Yong-cheng, "An improved TF-IDF approach for text classification", *Journal of Zhejiang University Science*, 2005.
- [21] M. Ghazi and D. Gangodkar, "Hadoop, MapReduce and HDFS: A Developers Perspective", *Procedia Computer Science*, vol. 48, 2015.
- [22] M. Nagao, "Natural Language Processing and Knowledge", *2005 International Conference on Natural Language Processing and Knowledge Engineering*.
- [23] "Hierarchical clustering", Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Hierarchical_clustering. [Accessed: 10- Jan- 2016].
- [24] "Complete-linkage clustering", en.wikipedia.org. [Online]. Available: https://en.wikipedia.org/wiki/Complete-linkage_clustering. [Accessed: 10- Jan- 2016].
- [25] "tweepy", github.com. [Online]. Available: <https://github.com/tweepy/tweepy>. [Accessed: 10- Feb- 2016].
- [26] "facebook4j", facebook4j.org. [Online]. Available: <http://facebook4j.org/en/index.html>. [Accessed: 10- Feb- 2016].
- [27] "hadoopy", hadoopy.readthedocs.org. [Online]. Available: <https://hadoopy.readthedocs.org/en/latest/>. [Accessed: 10- Feb- 2016].

Design Pattern for Multilingual Web System Development

Habes Alkhraisat

Abstract— Recently- Multilingual WEB Database system have brought into sharp focus the need for systems to store and manipulate text data efficiently in a suite of natural languages. While some means of storing and querying multilingual data are provided by all current database systems. In this paper, we present an approach for efficient development multilingual web database system with the use of object oriented design principle benefits. We propose functional, efficient, dynamic and flexible object oriented design pattern and database system architecture for making the performance of the database system to be language independent. Results from our initial implementation of the proposed methodology are encouraging indicating the value of proposed approach.

Index Terms— Database System, Design Pattern, Inheritance, Object Oriented, Structured Query Language.

I. INTRODUCTION

If your business or service addresses an international audience, having a multilingual system increases your ability to reach your target markets. The rapidly accelerating trend of globalization of businesses and the success of database system solutions require data to be stored and manipulated in many different natural languages. As the primary data repository for such applications, database systems need to be efficient with respect to multilingual data. While all current database systems support some means of storing and manipulating such data, there has been no object oriented pattern design in this regard.

Building a multi-language database system is not a trivial task and you will encounter many problems on this way, and one of them is how store and retrieve the content for each language. You may perform a small research on the Web and find enough resources about it, but there is no magic solution, each solution depends on personal requirements, size of the database, complexity of system, etc. Therefore, this paper propose a design pattern, as standard solution to a multilingual system programming, which enable large scale reuse and improve developer communication.

Design patterns provide a high-level language of programmers discourse to describe their systems and to discuss solutions to common problems. This language comprises the names of recognizable patterns and their elements. The proper and intelligent use of patterns will guide a developer into designing a system that conforms to well-established prior practices, without stifling innovation.

The remainder of this paper is organized as follows: Section 2 introduce the object-oriented design patterns. Section 3 explores a set of requirements to be supported by the translator design pattern with appropriate examples. . Section 4 gives development practices for translator design pattern.

II. OBJECT-ORIENTED DESIGN PATTERNS

Concept of Design patterns has been introduced by Christopher Alexander in civil architecture in 1977; they were later adapted to software design. Christopher Alexander says, "Each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million time over, without ever doing it the same way twice" [1].

Design patterns are solutions to software design problems you find repeatedly in real-world application development. Patterns are about reusable designs and interactions of objects. In general, a pattern has four elements: name, problem, solution, and consequences [2]. The pattern name describes a design problem, its solutions, and consequences in a word or two. The problem describes when to apply the pattern and explains the problem context. Pattern solution describes the elements that make up the design, their relationships, responsibilities, and collaborations. The solution does not describe a particular concrete design or implementation, because a pattern is like a template that can be applied in many different situations.

The consequences are the results and trade-offs of applying the pattern. The consequences are critical for evaluating design alternatives, for understanding the costs, and benefits of applying the pattern. The consequences for software often concern space and time trade-offs. They may address language and implementation issues as well. Since reuse is often a factor in object-oriented design, the consequences of a pattern include its impact on a system's flexibility, extensibility, or portability.

III. PROPOSED TRANSLATOR DESIGN PATTERN

If your business or service addresses an international audience, having a multilingual website increases your ability to reach your audience. Rather than reproducing the web system in various languages, translator pattern has a quicker way to create a multilingual website that will appeal to your site visitors wherever they may be.

The translator pattern separates the system from how its objects is created, composed, and represented. It increases the system's flexibility in terms of the what, who, how, and when of object creation and the data translated. Translator pattern encapsulate the knowledge of how the data is translated, and which classes a system uses, but they hide the details of how the instances of these classes are created and put together.

A. Translator Design Pattern Role

The purpose of the Translator pattern is to generate the SQL statement to retrieve the data from the database based on the language setting selected by the web users. Moreover, the Translator pattern use the join operation between two relations, one relation contains all numerical value for a given problem, while the text data for that problem is stored in the second relation.

B. Translator Design Pattern Illustration

Any database system solutions require data to be stored and manipulated in many different natural languages would benefit from the translator pattern. As an example, consider online publishing system that is store and display content of article in many different natural languages. There are many ways to build such system, as putting article content in one relation (Figure 1-A). Alternatively, we suggest in Translator design pattern the structure in figure 1.B, where any relation is divided into two relations one for data unilingual, and the other for multilingual data. The beauty of this addition is that: unlimited number of languages, smart translation, and protection against duplication of content.

C. Translator Design Pattern Design

An important part of each pattern's description is a Unified Modelling Language (UML) class diagram. UML diagram shows the players in the pattern. Figure 2 shows UML diagram for the Translator pattern. The players in the pattern are:

- 1) Singleton design pattern to ensure that there is only one instance of a class, and that there is a global access point to that object. The pattern ensures that the class is instantiated only once and that all requests are directed to that one and only object. A special care should be taken in multithreading environments when multiple threads must access the same resources through the same singleton object [1].
- 2) Translator pattern provides the implementation for DML statement
- 3) Class and class translation

D. Translator Design Pattern Implementation

In this sub section, we will explain the implementation of proposed Translator pattern using PHP language. A good implementation of the Translator pattern in PHP relies on the precise interpretation of language rules regarding construction and inheritance. The simple theory code for the Translator pattern is shown in Example 1. As shown in the example the Translator pattern has been built with the use of Singleton design pattern. The benefit of Singleton pattern, it ensures that only one of a class can be built and that all users are directed to it [3, 4].

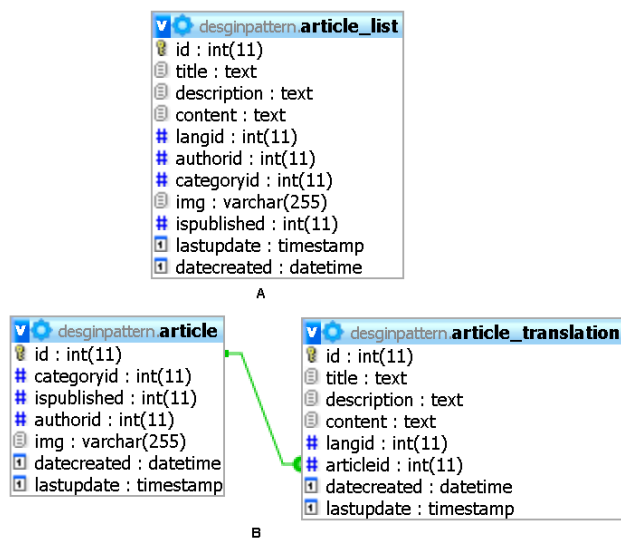


Fig. 1. Translator pattern illustration—(A) a article schema without applying the Translator Pattern, (B) new article schema for Translator Pattern

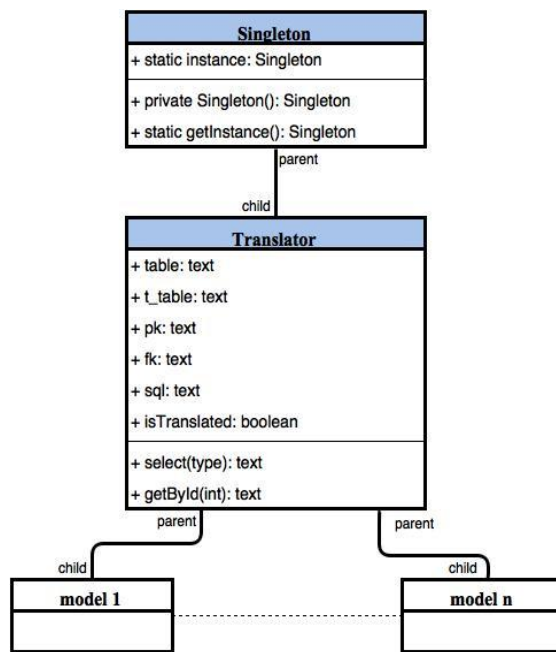


Fig. 2. Translator design pattern

Example 1. PHP implementation for Translator Design Pattern

```

class Singleton{
    private static $_singleton;
    private static $_connection;
    private $connectionInfo = array ();
    private function __construct () {
        //add the connection to database server
    }
    public static function getInstance () {
        if (is_null (self::$_singleton))
            self::$_singleton = new database ();
        return self::$_singleton;
    }
}
    
```

```

class translator extends Singleton{
    public $table;
    public $t_table;
    public $pk;
    public $fk;
    public $sql;
    public function __construct(){
        parent::getInstance();
        $this->t_table = $this-
>table.'_translation';
        $this->fk = $this->table.'_'. $this->pk;
    }
    public function select(){
        if (empty($this->isTranslated)||$this-
>isTranslated==FALSE){
            $this->sql = "select * from ".$this-
>table;
        }else{
            $this->sql = 'select ' . $this-
>table.'.*,';
            $this->sql.=$this->t_table.'.title,';
            $this->sql.=$this-
>t_table.'.description,';
            $this->sql.=$this->t_table.'.content
from';
            $this->sql.=$this->t_table.'.langid
from';
            $this->sql.=$this->table.' inner join
'.$this->t_table;
            $this->sql.=' on '.$this-
>table.'.id='.$this->fk;
            $this->sql.=' where ' ;
            $this->sql.=
'langid.='.$_SESSION["langid"];
        }
        return $this->sql;
    }
    public function getById($id){
        $this->sql = $this->select();
        $this->sql.=" and " . $this-
>table.".id=" . $id;
        return $this->sql;
    }
}

```

The translator class maintains the following instance variables:

- **table**: the name of translatable table, which contains mostly numerical data,
- **t_table**: the name of translatable table translations, which offers content in several languages. The **t_table** attribute can be constructed from the main table by the translator constructor (line-23), just adds the suffix "_translation" at the table name
- **pk**: the primary key of the table,
- **fk**: foreign key in the translation table that refers to the primary key of table, . The **fk** attribute can be populated by the translator constructor.
- **sql**: the SQL query generated by the translator pattern.

The translator constructor creates a connection to the database server by calling the static *getInstance* method.

The *select* method of translator generates the SQL statement by joining the translatable table, and of translatable table translations to retrieve the data based on the language that the user selected. In this implementation for the Translator pattern, the user-selected language is stored in session.

IV. DEVELOPMENT PRACTICES

We have already looked at some theory code and discussed the Translator pattern concept. It this section illustrates by example how to use of proposed pattern in multilingual web system development. The task for the example is let us say we wanted to develop a system for online publishing system using Translator pattern, as described in the "Illustration and Implementation" section, earlier.

A. Database Design

For unlimited number of languages, smart translation, protection against duplication of content, and the efficient use of translator pattern to build a publishing system, we suggest the following database schema. Where any relation is divided into two relations one for data unilingual, and the other for multilingual data. As example the language relation is divided into 2 relations, the first is language to store data unilingual, and language translation relation is used to store multilingual data (Table 1 and 2). Moreover, Table 3, 4, 5, and 6 show another example for storing genres of article and the article content.

B. Presenting Language Options

First, you will need to choose the primary language for your system. Primary language is the language the majority of content is written in.

A multilingual system is useless without the ability to change languages. Often times you will find multilingual system use a dropdown; placed top on the page. You might also find switchers in the footer. Whichever pattern you go for, make sure that the dropdown is easy to see and access.

In the suggested pattern, the selected language is stored in a session variable called *langid*. For example, if the selected language is English the value for variable *\$_SESSION["langid"]* is 1, and if it is Arabic the value for *\$_SESSION["langid"]* is 2; where 1 and 2 are the id of language (Table 1 and 2).

C. Load the content

In this section, we will give an examples demonstrating Translator pattern and describe how to use the Translator pattern to develop multilingual web system. The main idea behind this first example: retrieve the list of all articles' genres based on specific language. To build the sub system for this example, we can observe two classes namely *genres* and *translator*, as shown below. The class *genres* is derived from class translator, and the class genres inherits all the member variables (*table*; *t_table*; *pk*; *fk*; *sql*) and methods (*select()*, and *getById(\$id)*) from its superclass translator. It further defines a variable called *isTranslated*, and its own constructors, as shown:

```

1 class genres extends translator {
2     public $isTranslated;
3     function __construct(){
4         $this->tbl = "genres";
5         $this->pk="id";
6         $this->isTranslated = true;
7         parent::__construct();

```

```

8     }
9     }
10    $obj = new genres();
11    $obj->select();
12    $obj->getById(1);

```

Line 4-6 initialize the member variables - the translatable table, and the primary key of translatable table. The member variable *isTranslated* is a flag indicates that the genres content in several languages.

Line 7 invokes the constructor of translator that does the following:

- creates an object from the singleton pattern – get a connection to the database,
- initializes instance variables *t_table* to the name of translatable table translations, in this case *t_table* is initialized to the values “*genres_translation*”,
- Initializes instance variables *fk* which represents the relationship between translatable table and translatable table translations, in this case the *fk* is initialized to “*genres-id*”.

Line 10 constructs instances *\$obj* of the class genres, and line 11 invokes the method *select()* belonging to a class genres. Suppose that the user selected-languages is English (i.e. *\$_SESSION["langid"]=2*), the translator pattern generate the following SQL statement:

```

Select genres.*, genres_translation.title,
genres_translation.description,
genres_translation.content,
genres_translation.langid from genres inner join
genres_translation on genres.id =
genres_translation.genres id where langid =1

```

Table 7 shows the result set of executing the generated query. As shown in the table, the result set is the content of genres is in English language. If the select language is Arabic, the Translator pattern generate the following SQL statement:

```

select genres.*, genres_translation.title,
genres_translation.description,
genres_translation.content,
genres_translation.langid from genres inner join
genres_translation on genres.id =
genres_translation.genres id where langid =2

```

Table 8 shows the result set of executing the generated query. As shown in the table, the result set is the content of genres is in Arabic language.

For retrieve specific tuple (Line 12) the method *getById(1)* belonging to a class genres is invoked. Assuming the selected language is English, the translator pattern generate the following SQL statement:

```

select genres.*, genres_translation.title,
genres_translation.description,
genres_translation.content ,
genres_translation.langid from genres inner join
genres_translation on genres.id =
genres_translation.genres_id where langid =1 and
genres.id =1

```

Table 9 shows the result set of executing the generated query. . If the select language is Arabic, the Translator pattern generate the following SQL statement:

```

selectgenres.*, genres_translation.title,
genres_translation.description,
genres_translation.content ,
genres_translation.langid from genres inner join
genres_translation on genres.id =
genres_translation.genres_id where langid =2 and
genres.id = 1

```

Table 10 shows the result set of executing the generated query. As shown in the table, the result set is the content of genres is in Arabic language.

The following shows another example to how to build article sub system using translator patter.

```

class article extends translator {
1  public $isTranslated;
2  function __construct() {
3      $this->tbl = "article";
4      $this->pk="id";
5      $this->isTranslated = true;
6      parent::__construct();
7  }
8  }
9

```

V. CONCLUSIONS AND FUTURE WORK

This paper has introduced and described translator design patterns for the development of multilingual web system. We have suggested new design pattern for a multi-lingual system where the content is written in more than one language. The information displayed in different languages is often the same, but maybe tailored for different audiences. Translator design patterns are mostly seen as solutions to a multi-lingual software design issues. The proposed translator pattern has the following advantages: unlimited number of languages, smart translation, and protection against duplication of content.

REFERENCES

- [1] C. Alexander, “A Pattern Language”, New York: Oxford University Press, 1977.
- [2] E. Gamma, R. Helm, R.Johnson, J. Vlissides, “Design Patterns: Elements of Reusable Object-Oriented Software”, 1st ed. Addison-Wesley Professional, 1994.
- [3] J. Bishop, “C# 3.0 Design Patterns”, 1st ed. O’Reilly Media, 2008.
- [4] A. Shalloway, J. Trott, J., “Design Patterns Explained: A New Perspective on Object Oriented Design”. 2nd ed. Addison-Wesley, 2004



Habes Alkhraisat was born in Amman, Jordan, in 1979. He received the B.S. in information technology from Al-Balqa Applied University in 2001, and M.S. degrees in computer science from the University of Jordan, in 2003 and the Ph.D. degree in System Analysis, Information Processing and Control from Saint Petersburg Electro Technical University (LETE), Saint Petersburg, Russia, in 2008.

Since 2009, he has been an Assistant Professor with the Computer Science department, Applied University. His research interest includes the system modeling, image processing, and biometric.

TABLE 9
LANGUAGE

Id	Code	Dir	Ispublished	Unicode	Datecreated	Lastupdate
1	En	Ltr	1		01-01-2016	00-00-0000
2	Ar	Rtl	1		01-01-2016	00-00-

TABLE 8
LANGUAGE-TRANSLATION

Id	title	description	Content	Language_id	langid	Datecreated	Lastupdate
1	English	Null	Null	1	1	01-01-2016	00-00-0000
2	Arabic	Null	Null	2	1	01-01-2016	00-00-0000

TABLE 7 GENRES

Id	Ispublished	parentid	Datecreated	Lastupdate
1	1	Null	01-01-2016	00-00-0000
2	1	1	01-01-2016	00-00-0000
0	0	1	01-01-2016	00-00-

TABLE 2
GENRES-TRANSLATION

Id	title	description	Content	genres_id	langid	Datecreated	Lastupdate
1	Computers	Computers	Computers	1	1	01-01-2016	00-00-0000
2	الكمبيوتر	الكمبيوتر	الكمبيوتر	1	2	01-01-2016	00-00-0000
3	Web	Web	Web	2	1	01-01-2016	00-00-0000
4	ويب	ويب	ويب	2	2	01-01-2016	00-00-

TABLE 1
ARTICLE

Id	genresid	ispublished	authorid	image-path	datecreated	lastupdate
1	2	1	1	/images/1/img1.jpg	01-01-2016	01-01-2016
2	3	1	1	/images/2/img2.jpg	01-01-	01-01-

TABLE 6
ARTICLE-TRANSLATION

Id	title	description	content	articleid	langid	datecreated	lastupdate
1	website design	website design	The term web design is normally used to describe the design process relating to the front-end (client side) design of a website including writing mark up. Web design partially overlaps web engineering in the broader scope of web development. Web designers are expected to have an awareness of usability and if their role involves creating mark-up then they are also expected to be up to date with web accessibility guidelines.	1	1	01-01-2016	01-01-2016
2	تصميم المواقع	تصميم المواقع	تصميم المواقع: هو عملية تخطيط وتنفيذ محتويات متعددة الوسائط عبر الشبكة (الإنترنت)، بواسطة أنماط التقنيات كلغات التوصيف المناسبة للعرض على متصفحات الإنترنت أو بنية واجهات المستخدم المبنية في الإنترنت.	1	2	01-01-2016	01-01-2016

TABLE 3
LIST OF ALL GENRES IN ENGLISH LANGUAGE

Id	Ispublished	parentid	Datecreated	Lastupdate	title	description	Content	langid
1	1	Null	01-01-2016	00-00-0000	Computers	Computers	Computers	1
2	1	1	01-01-2016	00-00-0000	Web	Web	Web	1

TABLE 4
LIST OF ALL GENRES IN ARABIC LANGUAGE

ID	Ispublished	parentid	Datecreated	Lastupdate	title	description	Content	langid
1	1	Null	01-01-2016	00-00-0000	الكمبيوتر	الكمبيوتر	الكمبيوتر	1
2	1	1	01-01-2016	00-00-0000	ويب	ويب	ويب	1
0	0	1	01-01-2016	00-00-0000	قواعد بيانات	قواعد بيانات	قواعد بيانات	1

TABLE 5
GET THE GENRES WITH ID =1 IN ENGLISH LANGUAGE

ID	Ispublished	parentid	Datecreated	Lastupdate	title	description	Content	langid
1	1	Null	01-01-2016	00-00-0000	Computers	Computers	Computers	1

TABLE 10
GET THE GENRES WITH ID =1 IN ARABIC LANGUAGE

ID	Ispublished	parentid	Datecreated	Lastupdate	title	description	Content	langid
1	1	Null	01-01-2016	00-00-0000	الكمبيوتر	الكمبيوتر	الكمبيوتر	1

A Model for Deriving Matching Threshold in Fingerprint-based Identity Verification System

Omolade Ariyo. O.

Research scholar

Department of Computer Science

University of Ilorin, Ilorin, Nigeria

Fatai Olawale. W.

Research Scholar

Department of Computer Science

University of Ilorin, Ilorin, Nigeria

Abstract

Currently there is a variety of designs and implementation of biometric especially fingerprint. There is currently a standard used for determining matching threshold, which allows vendors to skew their test results in their favour by using assumed figure between -1 to +1 or values between 1 and 100%. The research contribution in this research work is to formulate an equation to determine the threshold against which the minutia matching score will be compare using the features set of the finger itself which is devoid of assumptions. Based on the results of this research, it shows that the proposed design and development of a fingerprint-based identity verification system can be achieved without riding on assumptions. Thereby, eliminating the false rate of Acceptance and reduce false rate of rejection as a result of the threshold computation using the features of the enrolled finger. Further research can be carried out in the area of comparing matching result generated from the threshold assumption with the threshold computation formulated in this thesis.

Keywords: Biometrics; Threshold; Matching; Algorithm; Scoring.

I. Introduction

Security issues seem to be one of the most important problems of contemporary computer science. One of the most

important branches of security is identification of users. Identification may be required for access control to buildings (e.g. Library), rooms, devices or information. In case of computer systems we say about access to software and data. The basic aim of identification is to make it impossible for unauthorized persons to access to the specified resources [1]. There are generally three solutions for performing secure identification:

- Token methods (something you have),
- Memory methods (something you know).
- Biometric methods (somebody you are).

The token method has two significant drawbacks. Firstly, the token may be lost or stolen. A person who finds or steals a token may have an access to all the resources that the proper owner of the token was able to access, and there is no possibility to find out if they are the person they claim to be. Secondly, the token may be copied. The easiness of making a copy is of course different for different kinds of tokens, but it is always technically possible.

Memory based methods identify people by checking their knowledge. The most popular memory methods are of course different kinds of passwords. The main drawback of this kind of methods is the unconscious selectivity of human memory. People may do their best to remember a password but they cannot guarantee that the information will not be forgotten.

Similarly to the token method when a malicious user knows a password it is impossible to check if they are the person they claim to be. The problems with token and memory-based methods are the main cause of increasing interest in methods of identification based on biometric information of a person.

The challenges in terms of vulnerability associated with both token-based (e.g. library card) and knowledge-based (e.g. personal library number) authentication approach has paved way for biometric authentication as the most promising authentication technology where certain services have to be restricted to only authorized library users. The popularity enjoyed by fingerprint being the oldest biometric authentication approach makes it more suitable in this regard for easier installation and maintenance [2]. It has been a major problem for university of Ilorin to control access to resources in her library since it was practically difficult to screen out illegal users from gaining entrance into the library. For this reason among others, fingerprint technology was proposed as the automated access control technique in this context. This explains the motivation of the researcher in this study.

Biometrics refers to the automatic identification of a person based on his or her physiological or behavioral characteristics. It includes fingerprint, iris, facial and retinal.

Biometrics technologies are becoming the foundation of an extensive array of highly secure identification and personal verification solutions. Today, biometric is being spotlighted as the authentication method because of the need for reliable security.

Fingerprint authentication has been in use for the longest time and bears more advantages than other biometrics [3]. The authors concluded that fingerprinting is one of the best known and frequently used Biometrics. In 1924, Federal Bureau of Investigation n (FBI) is already known to have maintained more than 250 million civil files of fingerprints for the purpose of criminal investigation and the identification of unknown casualties (Federal Bureau of Investigation1908-2008). Now is being used in numerous field including financial, medical, e-

commerce and customer application as a secure and effective authentication method.

II. Research Objectives

Firstly, to design a model for threshold through which minutia matching can be determined finger-prints. Secondly, the main objective of this research work is to develop a Fingerprint based identity verification system for library users and

[5], concluded that in order to claim that two fingerprints are from the same finger, consider several factors which includes:

- i) Global pattern configuration agreement, which means that two fingerprints must be of the same type,
- ii) Qualitative concordance, which requires that the corresponding minutia details must be identical,
- iii) Quantitative factor, which specifies that at least a certain number (a minimum of 12 according to the forensic guidelines in the United States) of corresponding minutia details must be found.
- iv) Corresponding minutia details, which must be identically inter-related. In practice, complex protocols have been defined for manual fingerprint matching and a detailed flowchart is available to guide fingerprint examiners in manually performing fingerprint matching.

III. Methodology

The proposed Fingerprint-based identity verification system, uses fingerprint verification. In verification, the system recognizes an individual by comparing his /her biometrics with a single-specified record in the database, making it a 1:1 verification mapping. In general, biometric verification consists of two (2) stages:

- Enrollment and

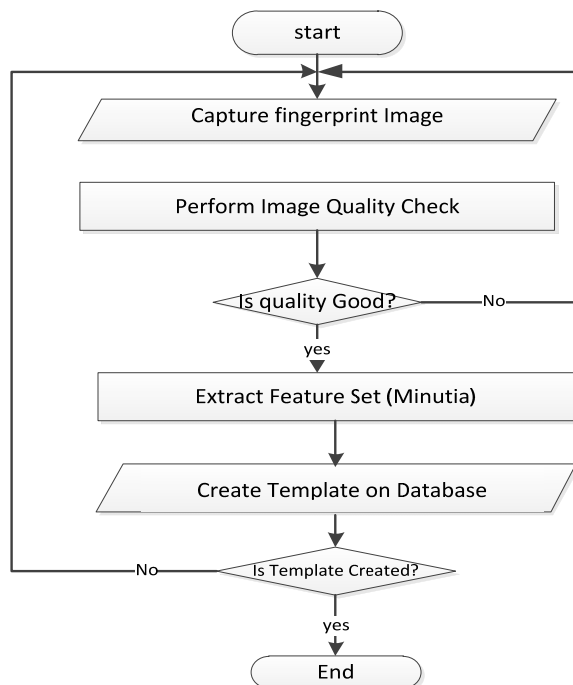
- Verification

A. Enrolment stage

During enrollment stage, the biometrics (fingerprint) of the Library user (Student, Lecturer and Registered Alumni) is captured by the administrator using a fingerprint reader and the unique features like minutia are extracted and stored in the database along with attributes which might not be physiological for later use.

Therefore, biometric characteristic of a subject is first captured by a biometric scanner to produce a sample. A quality check is often performed to ensure that the acquired sample can be reliably processed by successive stages. A feature extraction module is then used to produce a feature set. The template creation module uses the feature set to produce an enrollment template. The proposed system then takes the enrollment template and stores it in the systems' Database storage as shown using Algorithm in (figure 1.0), together with other information about the user (such as an identifier, name, gender, Matric number, level etc.).

- Figure 1.0. Flowchart for the Enrollment Module.



B. Verification stage

- During verification stage, the fingerprint of the library user is re-captured again, the image is also checked for quality and the feature sets are extracted and are then compared (using the matching Algorithm) with the ones already stored as template in the database in order to determine a match. Verification is typically used for positive recognition, where the aim is to prevent multiple people from using the same identity. Fingerprint verification is to verify the authenticity of one person by his fingerprint. There is one-to-one comparison in this case. For the purpose of this thesis a new matching algorithm is proposed that will modify the Minutia-based matching for better performance.

Furthermore, verification process is responsible for confirming the claim of identity of the subject. During the recognition phase, an identifier of the subject (such as Matric Number or Library ID) is provided (e.g., through a keyboard) to claim an identity; the biometric scanner captures the characteristic of the user and converts it to a sample, which is further processed by the feature extraction module to produce a feature set (Minutia). The resulting feature set is fed to the matcher, where it is compared against the claimed enrolled template of the user (retrieved from the system storage based on the user's identifier). The verification process produces a match/non-match (Verified/Non-Verified) decision by comparing the score and the threshold values of the fingerprint image.

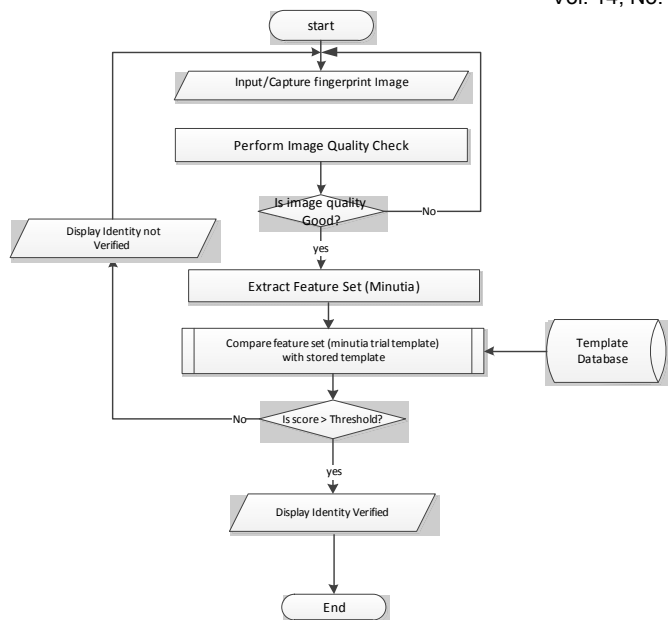


Figure 2.0. Flowchart for the Verification Module.

IV. The Proposed Matching Algorithm for the System.

Fingerprint features can be analyzed at both global and local level. When analyzed at the global level, the fingerprint pattern exhibits one or more regions where the ridge lines assume distinctive shapes (characterized by high curvature, frequent termination, etc.). These regions (called singularities or singular regions) may be classified into three typologies: loop, delta, and whorl.

Singular regions belonging to loop, delta, and whorl types are typically characterized by \cap , Δ , and O shapes, respectively [6]. Several fingerprint matching algorithms pre-align fingerprint images according to a landmark or a centre point, called the core. The core corresponds to the centre of the north most loop type singularity. For fingerprints that do not contain loop or whorl singularities, it is difficult to define the core. In these cases, the core is usually associated with the point of maximum ridge line curvature. Unfortunately, due to the high variability of fingerprint patterns, it is difficult to reliably locate a registration (core) point in all the fingerprint images [6].

Minutiae-based matching is basically a point pattern matching problem that is generally intractable because it

encounters the minutiae correspondence problem. It can be quite difficult to obtain the minutiae correspondence because the new image can be subject to transformation such as rotation, translation or even deformation. The location and direction errors of the detected minutiae as well as presence of spurious minutiae or absence of genuine minutiae can cause a lot of incongruity in the minutiae correspondence.

Let S and L represent the stored template and the live template respectively in an FIVS (Fingerprint-based Identity Verification System). If we consider S and L as feature vectors then each minutia is an element of the feature vector. Each minutia can be described by a number of attributes such as its location on the image, orientation and the type. Most common minutiae matching algorithms consider each minutia as a triplet $m = \{x, y, \theta\}$ that indicates the x, y minutia location coordinates and the minutiae angle θ [5].

$$S = \{m_1, m_2, \dots, m_m\}, \quad m_i = \{x_i, y_i, \theta_i\}, \quad I = 1 \dots m \quad (1)$$

$$L = \{m'_1, m'_2, \dots, m'_n\}, \quad m'_j = \{x'_j, y'_j, \theta'_j\}, \quad J = 1 \dots n \quad (2)$$

Where m and n denote the number of minutiae in S and L respectively.

A minutia m'_j in L and a minutia m_i in S are said to be 'matching' if the spatial distance (sd) between them is smaller than a given tolerance r_0 and the direction difference (dd) between them is smaller than an angular tolerance θ_0 see the figure below:

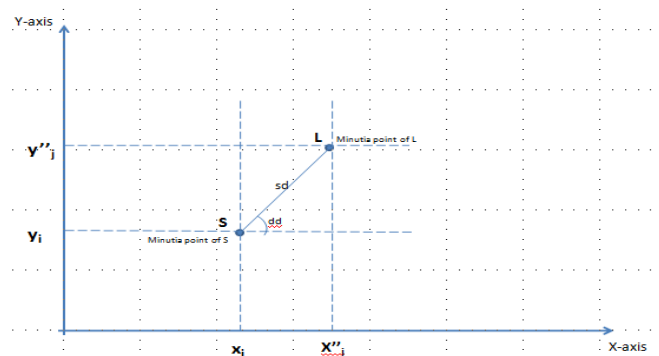


Figure3.0: showing the two minutia S and L on a plain.

$$sd(m'_j, m_i) = \sqrt{(x'_j - x_i)^2 + (y'_j - y_i)^2} \leq r_0, \text{ and} \quad (3)$$

$$dd(m'_j, m_i) = \min(|\theta'_j - \theta_i|, 360^\circ - |\theta'_j - \theta_i|) \leq \theta_0 \quad (4)$$

(6)

Equation (4) takes the minimum of $|\theta'_j - \theta_i|$, and $360^\circ - |\theta'_j - \theta_i|$ because of the circulatory nature of the angles. The tolerance boxes (or hyper-spheres) defined by r_0 and θ_0 are required to accommodate the unavoidable errors made by feature extraction algorithms and to count for the small displacement that cause the minutiae position to change.

In many algorithms, alignment of the stored and query templates is mandatory to maximize the number of matching minutiae in terms of their corresponding position and orientation. When two fingerprints are correctly aligned, the displacement (in x and y) and rotation (θ) are recovered and it likely to compensate other geometrical transformations:

- If two fingerprint images have been taken by scanners operating at different resolutions then scaling needs to be done.
- Other distortion-tolerant geometrical transformations could be useful to match minutiae in case one or both of the fingerprints is affected by severe distortions.

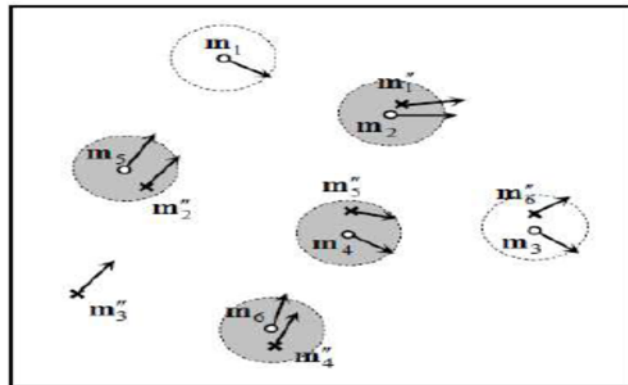
In designing a matching algorithm, the tolerance box should be carefully calculated as adjustment for any other geometrical transformations beyond translation and rotation may results in additional degrees of freedom to the minutiae matcher thus lead to a huge number of new possible alignments which significantly increases the chance of incorrectly matching two fingerprints from different fingers.

If $map()$ is a function that maps a minutia m'_2 (from L) into m''_j according to a given geometrical transformation; for example, by considering a displacement of $[\Delta x, \Delta y]$ and a anticlockwise rotation θ around the origin:

$$map_{\Delta x, \Delta y, \theta}(m'_j = \{x'_j, y'_j, \theta'_j\}, m''_j = \{x''_j, y''_j, \theta'_j + \theta\}), \quad (5)$$

$$where \begin{bmatrix} x''_j \\ y''_j \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x'_j \\ y'_j \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

Let $mm()$ is an indicator function that returns 1 in the case where the



minutiae m''_j and m'_i match according to equations (3) and (4)

$$mm(m''_j, m'_i) = \begin{cases} 1 & sd(m''_j, m'_i) \leq r_0 \text{ and } dd(m''_j, m'_i) \leq \theta_0 \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Then the matching problem can be formulated as

$$maximise_{\Delta x, \Delta y, \theta, P} \sum_{i=1}^m mm(map_{\Delta x, \Delta y, \theta}(m'_{P(i)}), m_i), \quad (8)$$

Where $P(i)$ is a function that determines the pairing between L and S minutiae; in particular, each minutia has either exactly one mate in the other fingerprint or has no mate at all.

The constraints are

1. $P(i)=j$ indicates that the mate of the m_i in S is the minutia m'_j in L.
2. $P(i)=null$ indicates that minutia m_i in S has no mate in L.
3. A minutia m'_j in L has no mate in S if $P(i) \neq j \quad \forall i=1, \dots, m$.
4. $\forall i=1, \dots, m, k=1, \dots, m, i \neq k \Rightarrow P(i) \neq P(k) \text{ or } P(i) = P(k) = null$ (this means that each minutia in L is associated with a maximum of one minutia in S, that is P is a bijective function).

Expression (8) requires that the number of minutiae mates be maximized, independently of how strict these mates are; in other words, if two minutiae comply with Equations (3) and (4), then their contribution to expression (8) is made independently of their spatial distance and of their direction difference.

Alternatives to expression (8) may be introduced where the residual (i.e., the spatial distance and the direction difference between minutiae) for the optimal alignment is also taken into account.

To comply with constraint (4) above, each minutia already mated has to be marked, to avoid mating it twice or more. Figure 5.0 shows an example of minutiae pairing given a fingerprint alignment.

Figure 4.0: Minutiae of L mapped into S coordinates for a given alignment. Minutiae of S are denoted by os , whereas minutiae of L are denoted by xs . Note that L minutiae are referred to as m'' , because what is shown in the figure is their mapping into coordinates of S. Pairing is performed according to the minimum distance. The dashed circles indicate the maximum spatial distance. The gray circles denote successfully mated minutiae; minutiae m_1 of S and minutiae m''_3 of L have no mates. Minutiae m_3 and m''_6 cannot be matched due to their large orientation difference [7].

To achieve the maximum pairing, a slightly more complicated scheme was adopted: in fact, incase when a minutia of L falls within the tolerance box of more than one minutia of S, the optimum assignment is that which maximizes the number of mates (see Figure 3.3).

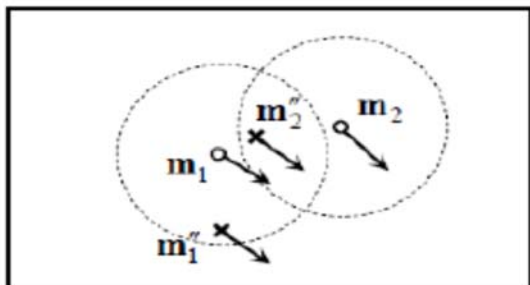


Figure 5.0: In this example, if m_1 was mated with m''_2 (the closest minutiae), m_2 would remain unmated; however, pairing m_1 with m''_1 , allows m_2 to be mated with m''_2 , thus maximizing equation(8) [7].

V. Similarity score

Unlike in manual matching performed by forensic experts where the number of matching minutiae is itself the main output of the comparison, automatic matching systems must convert this number into a similarity score Q . This is often performed by simply normalizing the number of matching minutiae (here

denoted by k) by the average number $(m + n)/2$ of minutiae set in S and L [7].

$$\text{score} = \frac{2k}{m+n} \tag{9}$$

However, further information can be exploited, especially in case of noisy images and limited overlap between S and L, to compute a more reliable score; in fact:

- Minutiae quality can be used to weight differently reliable and unreliable minutiae pairs: the contribution from a pair of reliable minutiae should be higher than that from a pair where at least one the two minutiae are of low quality [8]. The quality of a minutia (and of a minutia pair) can be defined according to the fingerprint quality in the region where the minutia lies and/or by keeping into account other local information.
- The normalization in Equation (9) tends to excessively penalize fingerprint pairs with partial overlap; a more effective normalization considers the number of minutiae belonging to the intersection of the two fingerprints after the optimal alignment have been determined.

VI. Matching Algorithm derived from the formulated problem.

For each pair of the minutia in S and L, find the translation and rotation parameters between the ridge associated with input minutia and the ridge associated with template minutiae and align the two minutiae patterns according to the estimated parameters and also the P the total number of minutia pair. Convert the template S and the input L patterns into the polar coordinate representations with respect to the corresponding minutia on which alignment is achieved and represent them as two symbolic strings by concatenating each minutia in an increasing order of radial angles:

$S = \{m_1, m_2, \dots, m_m\}$, $m_i = \{x_i, y_i, \theta_i\}$, $I = 1 \dots m$

$L = \{m'_1, m'_2, \dots, m'_n\}$, $m'_j = \{x_j, y_j, \theta_j\}$, $J = 1 \dots n$

where m and n denote the number of minutiae in S and L respectively while x, y minutia location coordinates and the θ minutiae angle.

Determine the tolerances both r_0 and θ_0 for spatial distance (sd) and direction difference (dd) respectively

Using the equations below determine the spatial distance (sd) and the direction difference (dd).

$$sd(m'_j, m_i) = \sqrt{(x'_j - x_i)^2 + (y'_j - y_i)^2} \leq r_0, \text{ and}$$

$$dd(m'_j, m_i) = \min(|\theta'_j - \theta_i|, 360^\circ - |\theta'_j - \theta_i|) \leq \theta_0$$

if the $sd \leq r_0$ and $dd \leq \theta_0$ For a minutia m'_j in S and m_i in L return 1, then there is a match at that point x, y .
else return 0 no match.

$$mm(m'_j, m_i) = \begin{cases} 1 & \text{if } sd(m'_j, m_i) \leq r_0 \text{ and } dd(m'_j, m_i) \leq \theta_0 \\ 0 & \text{otherwise.} \end{cases}$$

count step 5 for all $i = 1 \dots m$ and $j = 1 \dots n$ and store count in k , repeat steps 4 and 5 until $I, j = m, n$ respectively.

Determine the matching similarity score (Q) using the formula:

$$score = \frac{k}{(n+m)/2}$$

Where K = number of matching minutia, while m, n are the number of minutia in S and L respectively.

Compute the threshold (T) that the score will be compare against, using the expression:

$$T = Q - (C / (n+m)) * 0.5$$

Where C is the total number of minutia that has no match and is given as $C = P - K$, P is the total minutia pair on the two

fingerprint (both match and non-match), K = number of matching minutia, while m, n are the number of minutia in S and L respectively, Q is the matching similarity score.

If $Q > T$ then there is a match on the fingerprint Else no-match.

End.

VII. Result

In this interface the admin selects the categories each library user belongs to.

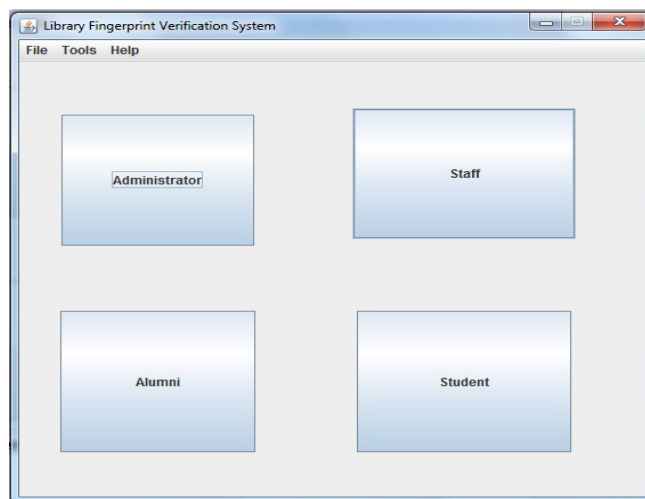


Figure 7.0: showing the landing page of the fingerprint system for Library users

A. The Enrollment Interface

In this window, user can enroll a fingerprint or match a fingerprint in database. User does not need to do pre-processing step by step (Note that before enroll/match, image must do image pre-processing).

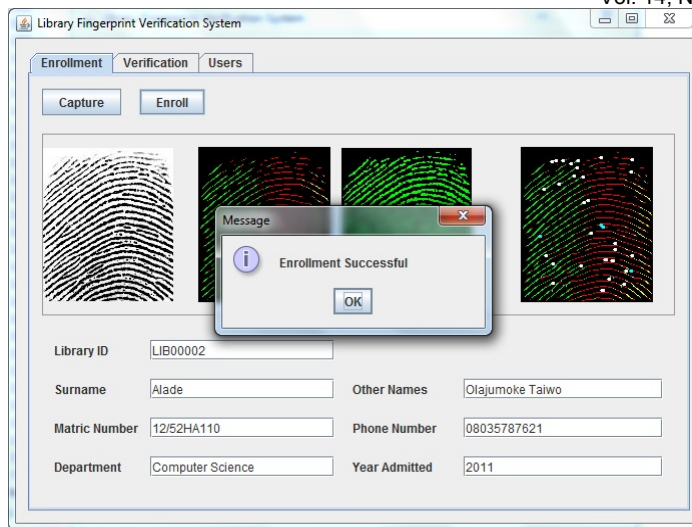


Figure 8.0: showing successful enrolment.

B. Verification Interface

Here the user lays claim on an identity, the system captures the user's fingerprint detail which is in turn verified if the claimed identity is right or not. It is as shown below.

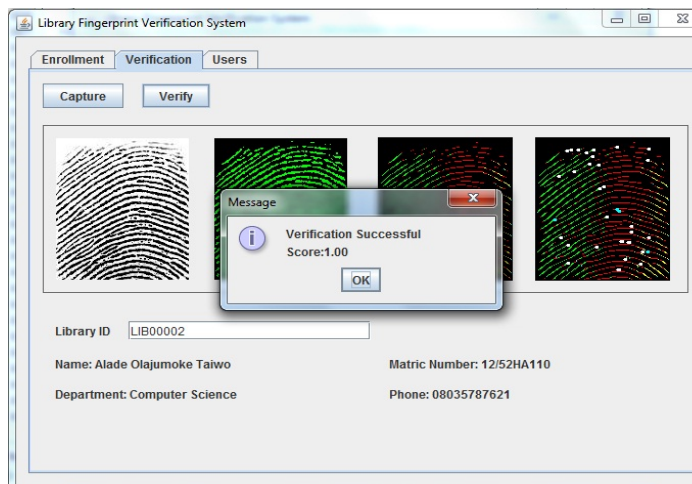


Figure 9.0: showing successful verification.

C. The Audit log

Every system should have an audit log in case of dispute resolution. The researcher has done justice to this. So that users activities can be monitored.

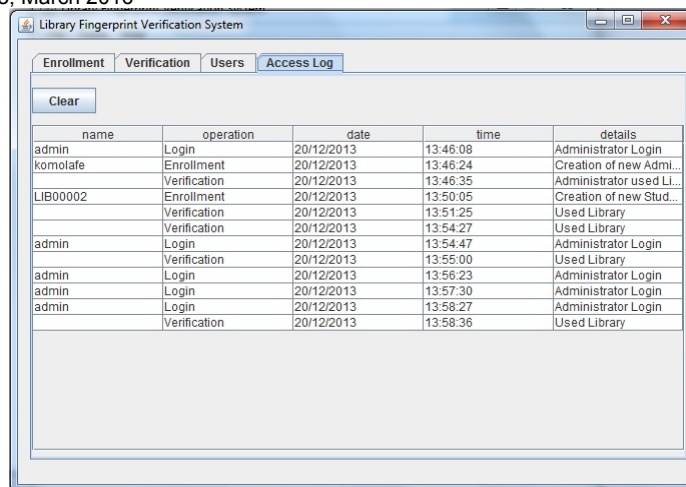


Figure 10.0: Showing the Audit log of the system.

VIII. Conclusion

Overall, the system was designed and implemented for Library users using the threshold model introduced in this research work, which seeks to use the minutia details of the fingerprint to determine the threshold against which a match or no-match result can be generated. Instead of the conventional threshold method where range of +1 to -1.

Further research can be carried out to test the usability of the system considering the threshold formulation introduced in this thesis for determining match and also compare it with existing matching system.

Secondly, a multimodal biometric approach to identity verification can be introduced to further enhance the security of access to the library

References

- [1] Paweł K. (2004). Human identification using eye movements. Doctoral Thesis. Silesian University of Technology. Retrieved May 22, 2013, from Mara University of Technology eprints repository.
- [2] Mark, B. S., Debjani, D., Vijay, K., & Ernst, B. (2008). A proposed study and analysis of user perceptions of Biometric acceptance retrieved March 1, 2011 from <http://www.decisionsciences.org/proceedings/DSI2008/docs/396-2784.pdf>.
- [3] Josphineleela R. and Dr. M. Ramakrishnan (2012), An Efficient Automatic Attendance System Using

Fingerprint Reconstruction Technique: (IJCSIS)
International Journal of Computer Science and
Information Security, Vol. 10, No. 3.

Computer Science. Fatai is a specialist in Network Design and
Implementation and also a Database Designer. He can be
reached through the following contacts. Email:
fataholawale@yahoo.com and phone number:
+2348032310132

- [4] Federal Bureau of Investigation (1908-2008). Retrieved May 15, 2013, from <http://issuu.com/faircountmedia/docs/fbi100>
- [5] Davide, M., Daria, M., Anil, K. J. & Salil, P., (2009). Handbook of Fingerprint Recognition, Springer-Verlag London Limited.
- [6] D. Maltoni (n.d), "A Tutorial on Fingerprint Recognition", Biometric Systems Laboratory - DEIS - University of Bologna, via Sacchi 3, 47023, Cesena (FC) – Italy.
- [7] K.M. Kryszczuk; P. Morier and A. Drygajlo, (2004). Study of the Distinctiveness of Level 2 and Level 3 Features in Fragmentary Fingerprint Comparison. in Proc. Workshop on Biometric Authentication (in ECCV 2004), LNCS 3087, pp. 124–133.
- [8] Yi Chen, (2009). Extended Feature Set and Touchless Imaging For Fingerprint Matching. Doctoral Thesis. Michigan State University. Retrieved July 22, 2013, from Michigan State University eprints repository.

Author's Profiles

Omolade Ariyo Olorunmeye, had a bachelor of science (B.Sc.) in Computer Science from the University of Ilorin, Ilorin, Nigerian and from the same school he bagged his Masters (M.Sc.) in Computer Science. He is a specialist in Software Engineering and also a skilled Project Manager. He is also already considering going for his PhD in Software Engineering.

Omolade can be reached through the following contacts.

Email: talk2tolade@yahoo.com and phone number:
+2347033377186

Fatai Olawale Waheed, had his Bachelor of Science (B.sc) in Computer science from the University of Ilorin, Ilorin, Nigeria and from the same institution he bagged his Masters (M.sc) in

A Sliding Mode Controller for Urea Plant.

M. M. Saafan*, M. M. Abdelsalam, M. S. Elksasy, S. F. Saraya, and F. F.G. Areed
Computers and Control Systems Engineering Department, Faculty of Engineering, Mansoura University,
Egypt.

ABSTRACT

The present paper introduces the mathematical model of urea plant and suggests two methods for designing special purpose controllers. The first proposed method is PID controller and the second is sliding mode controller (SMC). These controllers are applied for a multivariable nonlinear system as a Urea Reactor system. The main target of the designed controllers is to reduce the disturbance of NH₃ pump and CO₂ compressor in order to reduce the pollution effect in such chemical plant. Simulation results of the suggested PID controller are compared with that of the SMC controller. Comparative analysis proves the effectiveness of the suggested SMC controller than the PID controller according to disturbance minimization as well as dynamic response. Also, the paper presents the results of applying SMC, while maximizing the production of the urea by maximizing the NH₃ flow rate. This controller kept the reactor temperature, the reactor pressure, and NH₃/CO₂ ratio in the suitable operating range. Moreover, the suggested SMC when compared with other controllers in the literature shows great success in maximizing the production of urea.

Keywords: Sliding mode controller, PID controller, urea reactor, Process Control, Chemical Industry, Adaptive controller, Nonlinearity.

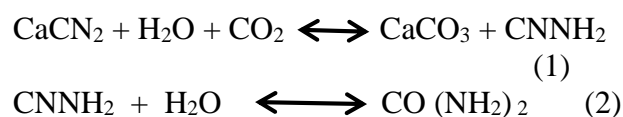
1. INTRODUCTION

Urea synthesis is one of the most important industries in agriculture as a fertilizer [1-2]. Urea is

generally used as a crude material in producing many chemical compounds including chemical solutions, plastics, adhesives and medical drugs. Urea is a natural compound in chemical equation $\text{CO}(\text{NH}_2)_2$.

In 1828, WOHLER discovered that urea can be produced from ammonia and cyanic acid in water solution. Since then, research on the urea preparation has continuously progressed [2-3].

In 1907, urea was produced on a limited industrial scale by dehydration of Cyanamid, which was obtained from calcium Cyanamid:



The first point for the industrial production of urea is the lab synthesis of BASAROFF [4], in which urea is obtained by dehydration of ammonium carbamate at increased pressure and temperature, the principle response happens in the fluid stage under pressure running from 13 to 25MPa. While, the reactor temperature is from 170 to 200°C [5-10].

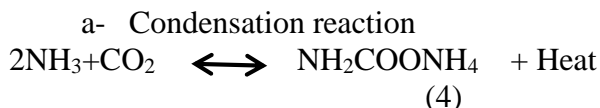
In the literature these are two suggested operation of reactors. The first operation was proposed by Frejacques [8] which is the originally seniority kinetic models for Urea reactor. The reaction is describes as shown in equ. (3).



In Equ (3), the reaction between NH₃ and CO₂ takes place under 145 bars pressure in order to form Urea. The urea is produced in one step reaction by reacting ammonia and carbon dioxide.

The second operation was generally accepted among researchers up to 1952. In this operation the urea is produced in two step reaction: first by reacting ammonia (NH₃) and carbon dioxide (CO₂), in gaseous phase, at high temperature (170 – 200°C), and high pressure (13 – 25 Mpa) (130 - 250 bars) in presence of ammonia carbamate (NH₂COONH₄) saturated solution in which the carbamate, in liquid phase is formed. This reaction is exothermic and fast. In the second reaction, which is endothermic and slow, the ammonium carbamate dehydrates to

produce water (H₂O) and urea (NH₂CONH₂) [10-15]. The reactions are described as:



b- Dehydration reaction



The reactor is composed such that its volume is sufficiently enormous for the wanted urea generation. The condensation reaction is the first reaction that takes place in urea plant. The condensation is depending on Pressure of condensation, reactor temperature, molar ratio between ammonia and carbon dioxide (NH₃/CO₂ ratio) and the molar ratio between water and carbon dioxide (H₂O/CO₂ ratio). The urea conversion increases when increasing temperature and NH₃/CO₂ ratio. To achieve the maximum urea conversion and reduce the pollution effect, the molar ratio between ammonia and carbon dioxide must be equal 3. The block diagram of urea synthesis process is shown in Fig.1. There are five stages for Urea synthesis plant. Synthesis stage, Recirculation stage, Desorption and Hydrolyzation Stage, Evaporation stage, and Granulation stage.

1- Synthesis Stage:

In which condensation reaction between NH₃ and CO₂ takes place under (140-146) bars pressure in order to form ammonium carbamate inside high pressure carbamate condenser (H.P.C.C). After that carbamate molecule loses water molecule and turns into urea molecule inside the reactor. The synthesis pressure consists of four equipment's working under high pressure about 145 bars High Pressure Carbamate Condenser (H.P.C.C), Reactor, High Pressure Stripper, and High Pressure Scrubber.

1- (H.P.C.C):

In which the first reaction between NH₃ and CO₂ both in gas phase takes place to form ammonium carbamate. This reaction is exothermic so in order to happen and continue heat produced from the reaction should be withdrawn. Heat is withdrawn by the production of low pressure steam about 4 bars. The equipment is a heat exchanger (shell and tube). The reaction happens inside the tubes and steam is produced in the shell side. Reaction between NH₃ and CO₂ gases doesn't happen in a complete form so a part of these gases should be left to go to the reactor in the gas phase to react inside it. Carbamate in the liquid and gas phases flow to the reactor.

2- Reactor

Reactor is a vertical cylindrical equipment. It contains a number of perforated sieve trays to allow gases to flow through perforations and liquid to flow through the distance between the wall and the tray. The liquid and gas phases' carbamate goes into reactor bottom out of the H.P.C.C and the gases reaction takes place first for the production of urea. Urea goes out of the reactor with about 35% concentration heading to the stripper. The reacted gases from reactor and inert gases goes to the scrubber.

3- H.P Scrubber

The outlet reactor gases is washed by the means of low concentration carbamate solution from recirculation stage for the absorption of NH₃ and CO₂.

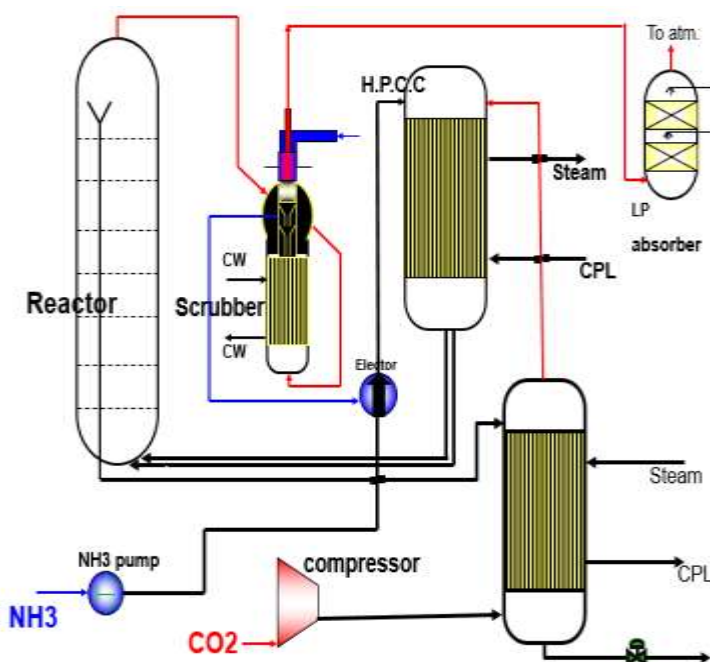


Fig.1. Block diagram of Urea synthesis process.

4- High Pressure Stripper:

Urea solution outlet reactor with concentration of about 35% goes to the stripper top. The rest is carbamate so it should be restored and directed again to the H.P.C.C. At the same time urea solution is concentrated from 35% to 57%. The stripping process is achieved by the counter current flow of CO₂ gas from its compressor.

Medium pressure steam (about 20 bars) is directed to the stripper shell side for the decomposition of carbamate and stripping it from urea solution using CO₂. Decomposed and stripped gases out solution of urea goes out of the stripper top with the CO₂ heading to the H.P.C.C. Urea solution comes out stripper bottom after it is concentrated to about 57% heading to recirculation stage with low pressure (about 4 bars) for completing urea concentration process.

2- Recirculation Stage:

After urea solution flows out of the stripper with concentration of about 57% the solution is flashed from 140 bars to 4 bars. In which the urea solution temperature is raised to get rid of NH₃ and CO₂ and increase urea solution to about 72%. Separated gases out of the solution head from rectifying column to the low pressure carbamate condenser (L.P.C.C).

3- Desorption and Hydrolyzation Stage:

The desorption idea is based on the stripping of NH₃ gas from diluted solution of about 5% NH₃-water using low pressure steam (about 4 bars). After ammonia gas is separated from ammonia water it flows to the reflux condenser where it is condensed and recycled into process again. The operation of hydrolyser is the same as the reactor but reversed. In which urea is hydrolyzed by means of 24 bars steam to its primary components NH₃ and CO₂. After that the outlet gases from the hydrolyser goes to the reflux condenser to be recycled into the process again.

4- Evaporation Stage:

In the evaporation stage urea solution is concentrated from about 72% to 96%. The process takes place in tow equipment's pre-evaporator and evaporator. Evaporation process takes place at 100°C

in the pre-evaporator and 132°C in the evaporator and under vacuum about 0.3 bar abs and the process happens under these circumstances to reduce urea hydrolysis and biuret formation.

5- Granulation Stage:

Urea fertilizer granules are produced through the injection of 96% concentration urea melt into the granulator on a bed of urea fine granules inside it. The basic idea is to fluidize the fertilizer with air and it can be briefed as follows:-

- 1- First urea solution is concentrated through evaporators and transformed from liquid to melt with 96% concentration.
- 2- Urea melt is injected on a bed of urea fine granules where the melt is pushed with pressurized air called atomization air through small nozzles, urea melt sprays out the nozzles to accumulate on the fine granules of the bed where urea granules are shaped in the desired size.

The flow chart of urea synthesis process is shown in Fig. 2.

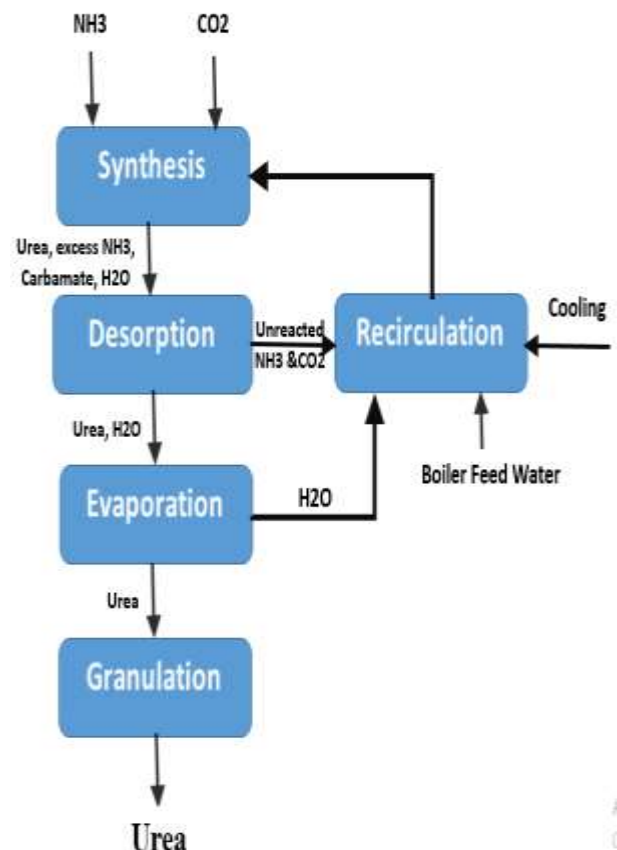


Fig.2. Flow chart of Urea synthesis process.

2. MATHEMATICAL MODEL OF UREA SYNTHESIS REACTOR

Equations (4) and (5) will be converted as ϵ_1 and ϵ_2 while the overall conversion as $\epsilon_1 * \epsilon_2 = \epsilon$:

$$\epsilon_1 = \frac{F_C + F_U}{F_{Ci} + F_{Di} + F_{Ui}} \quad (6)$$

$$\epsilon_2 = \frac{F_U}{F_C + F_U} \quad (7)$$

$$\epsilon = \epsilon_1 * \epsilon_2 = \frac{F_U}{F_{Ci} + F_{Di} + F_{Ui}} \quad (8)$$

Where, F_U , F_C and F_D are the flow rates of urea, carbamate and Carbene dioxide respectively. F_{Ci} , F_{Di} and F_{Ui} , are the initial flow rate of urea, carbamate and carbon dioxide respectively.

The flow rates of urea F_U , carbamate F_C , carbon dioxide F_D , ammonia F_N , water F_H and total rate F_T which is the sum of all individual flow rate are represented as:

$$F_U = \epsilon * (F_{Ci} + F_{Di} + F_{Ui}) \quad (9)$$

$$F_C = (\epsilon_1 - \epsilon) * (F_{Ci} + F_{Di} + F_{Ui}) \quad (10)$$

$$F_D = (1 - \epsilon_1) * (F_{Ci} + F_{Di} + F_{Ui}) \quad (11)$$

$$F_N = (a - 2\epsilon_1) * (F_{Ci} + F_{Di} + F_{Ui}) \quad (12)$$

$$F_H = (b + \epsilon) * (F_{Ci} + F_{Di} + F_{Ui}) \quad (13)$$

$$F_T = (1 + a + b + \epsilon - 2 * \epsilon_1) \quad (14)$$

Where a and b, are ammonia and water feed ratios which are represented as:

$$a = \frac{F_{Ni}}{F_{Ci} + F_{Di} + F_{Ui}} \quad (15)$$

$$b = \frac{F_{Hi}}{F_{Ci} + F_{Di} + F_{Ui}} \quad (16)$$

The rate of reactions for CO_2 , NH_3 and Urea are represented as:

$$r_{CO_2} = -k_{1F} * (Cd * Cn^2 - \frac{Cc}{k_1}) \quad (17)$$

$$r_{Carb} = k_{1F} * (Cd * Cn^2 - \frac{Cc}{k_1}) - k_{2F} * (Cc - Cu * \frac{Cw}{k_2}) \quad (18)$$

$$r_{Urea} = k_{2F} * (Cc - Cu * \frac{Cw}{k_2}) \quad (19)$$

Where, k_{1F} and k_{2F} are kinetic for the forward two urea reactions equations, Cd, Cc, Cu are molar flow rates for carbon dioxide, carbamate and urea, k_1 and k_2 are equilibrium constants.

$$k_i = k_0 * e^{\frac{-E_a}{R * T}} \quad (20)$$

Saturated vapor pressure is represented by Clausius-Clapeyron

$$\ln p_{NH_3}^s = -25.07T^{-1} + 56.321 \ln T - 0.2625T + 1.753 * 10^{-4}T^{-2} - 258.139 \quad (21)$$

$$\ln p_{CO_2}^s = -2370.26T^{-1} - 0.591 \ln T - 1.178 * 10^{-2}T + 1.598 * 10^{-5}T^2 + 15.272 \quad (22)$$

$$\ln p_{H_2O}^s = -5231.82T^{-1} - 6.167 * 10^{-2} \ln T - 3.291 * 10^{-3}T + 1.222 * 10^{-6}T^2 + 13.183 \quad (23)$$

The temperature must be in range 293 – 405 K in order to the equation (21) is true but in range 216 – 304 K for equation (22).

This paper presents the simulation results and analysis for the suggested control techniques for controlling the urea reactor system, applying SMC for increasing production, reducing the pollution effect and improving stability. In this paper, two design methods are suggested to minimize and reject

any disturbance applied to NH₃ Pump and CO₂ compressor. The two methods are suggested to achieve the maximum feeding flows of NH₃ and CO₂. This goal will be keeping the molar ratio between ammonia and carbon dioxide equal 3 to reduce the unreacted NH₃ and Co₂, and increasing the urea production, the temperature of reactor and pressure of reactor in operating ranges. The first method is based on PID controller techniques. We have two PID controllers to minimize the disturbance of NH₃ pump and CO₂ compressor. In the first PID controller, when the urea system is subjected to disturbance due to flow rate of NH₃, the reactor suffers from flow rate of NH₃ oscillations. The PID controller tries to minimize and reject any internal or external disturbance by adjusting the NH₃ pump. The second PID controller is used to minimized and reduce any disturbance due to flow rate of CO₂ by adjusting the CO₂ compressor. The second method is using SMC for NH₃ pump and CO₂ compressor. This paper arranged as the following. Section 2 represent the mathematical model of urea synthesis reactor. Section 3 describe the suggested method PID controller. Section 4 describe the suggested methods SMC. Section 5 describe how to using SMC to maximize the production and reducing the pollution effect by maximizing the feeding flows. Section 6 presents the simulation results obtained from the urea synthesis reactor. Then, a comparison analysis between two methods and then, a comparison analysis with previously related work.

3. A SUGGEST PID CONTROLLER FOR REDUCING SYSTEM POLLUTION.

The PID controller block diagram is shown in fig 3.1. PID controller used to remove any disturbances of NH₃ Pump and CO₂ compressor in Urea system. The controller uses the error signal *e* to produce a control signal *u* to control the flow through the NH₃ pump and CO₂ compressor [16-17]. The output control signal *u*(*t*) is described by

$$u(t) = K_p (e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt})$$

Or,

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt}$$

Where, K_p is the proportional gain, T_i is the integral time, T_d is the derivative time, $e(t)$ is the error signal, K_i is the integral gain, and K_d is derivative gain.

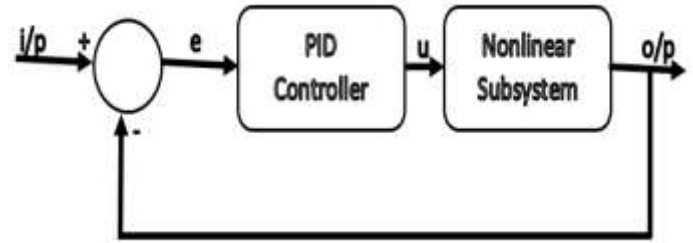


Fig.3.1. the PID block diagram.

4. A SUGGEST SLIDING MODE CONTROLLER FOR REDUCING SYSTEM POLLUTION.

The SMC has long known as a particularly suitable method for nonlinear systems with disturbances. The principle perfect of sliding mode control may be on implement the movement of the sliding mode on predefined switching surfaces in the framework state space utilizing spasmodic control. Those discontinues surfaces or those exchanging pronouncement if make chosen such-and-such sliding movement might show wanted Progress for movement done understanding with specific execution paradigm. Those strategies of the classical control theory, for example, eigenvalue placement alternately Linear-quadratic controller (LQR) to straight systems, camwood make relevant will pick fitting exchanging surfaces. Then, the discontinuous control needs to be chosen such that any states outside of the discontinuity surface are enforced to reach the surface in finite time. Accordingly, sliding mode happens along those surface, and the framework takes after those fancied framework flow [18]. The SMC block diagram is shown in fig.3.2.

To design the SMC, two stages are defined. The first stage is define a sliding surface of the process dynamics. The second stage is to design a feedback control law such any process's trajectory outside the sliding surface is driven to reach the surface is a finite time and keep on it. The control $u(x, t)$ with its respective entry $u_i(x, t)$ has the form

$$u_i = \begin{cases} u_i^+ & \text{if } s_i(x) > 0 \\ u_i^- & \text{if } s_i(x) < 0 \end{cases}$$

Where, $s_i(x)$ switching functions, u_i^+ and u_i^- are continuous functions. Since $s_i(x)=0$ is called a switching surface.

The state trajectory under the reaching condition is satisfied which is called reaching mode [19].

Existing of sliding mode:

Reaching Condition

$$\dot{s} < -\sigma \quad \text{sgn}(s), \quad \sigma > 0$$

Where, σ is positive constant.

Sliding condition

$$\lim_{s \rightarrow 0^+} \dot{s} < 0, \quad \lim_{s \rightarrow 0^-} \dot{s} > 0$$

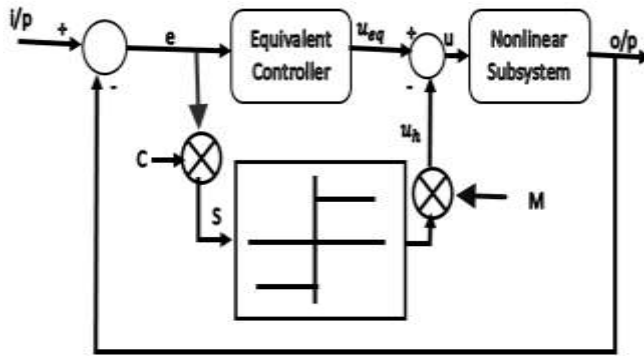


Fig.3.2. the block diagram of SMC.

The equivalent controller is responsible for generating the u_{eq} for controlling the NH3 pump and CO2 compressor. While u_h is the hitting control signal. The constant M is used to adjust the signal to be in range 0.5 to 1. The constant c is used for amplifying the error. So it should be more than unity, and not large for reducing the chattering effect [20-21].

5. A TECHNIQUE FOR INCREASING UREA PRODUCTION AND REDUCING POLLUTION EFFECT.

Nm3 SMC is used to increasing the urea production and reducing the pollution effect. This goal is achieved by maximizing the flow rates of NH3 and CO2 while keeping the NH3/CO2 ratio to reducing the unreacted NH3 and CO2 to reduce the pollution effect. Also, the reactor pressure and temperature keeping in operating range. We have three SMCs, the first one used to control the flow rate

of NH3, the second is used to control the CO2 flow rate to keep the optimal NH3/CO2 ratio, and the last one is used to reducing the reactor pressure and temperature disturbances to keeping in operating range. The urea production can be set by selecting the set point of the flow rate of NH3 controller. The NH3 flow rate multiply by a constant to produce the setpoint of CO2 flow rate. The function of CO2 flow rate controller is to keep the NH3/CO2 ratio by increased the CO2 flow rate. The amount of urea produced will be increased when augmented the setpoint of the NH3 flow rate controller.

6. SIMULATION RESULTES.

When the system is subjected to disturbances for NH3 Pump and CO2 compressor, the reactor suffers from NH3 and CO2 oscillations. The PID and SMCs tries to remove the disturbances in NH3 Pump and CO2 compressor. The molar ratio between ammonia and carbon dioxide must be equal three.

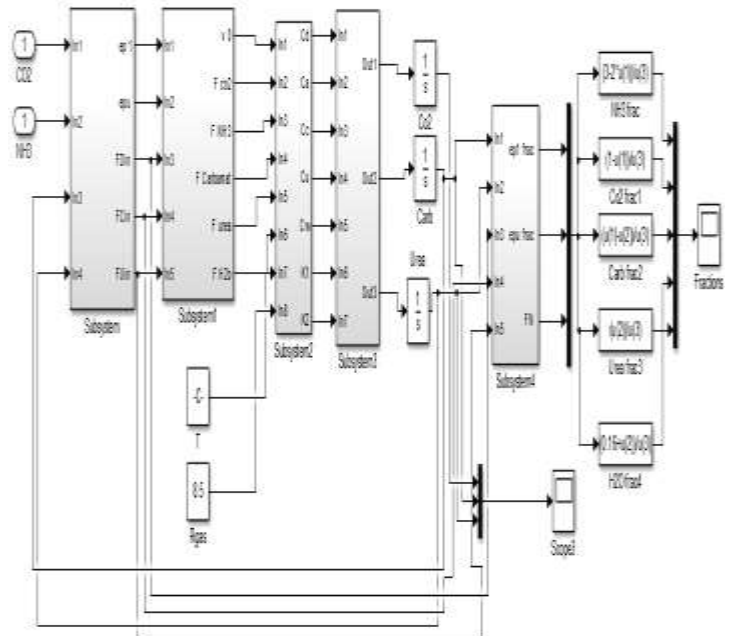


Fig.4. Simulink block diagram of the urea reactor model.

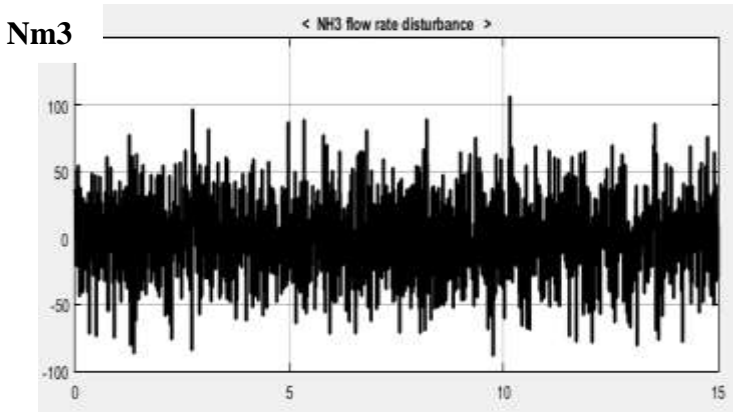


Fig.5. NH3 flow rate disturbance. T(Sec)

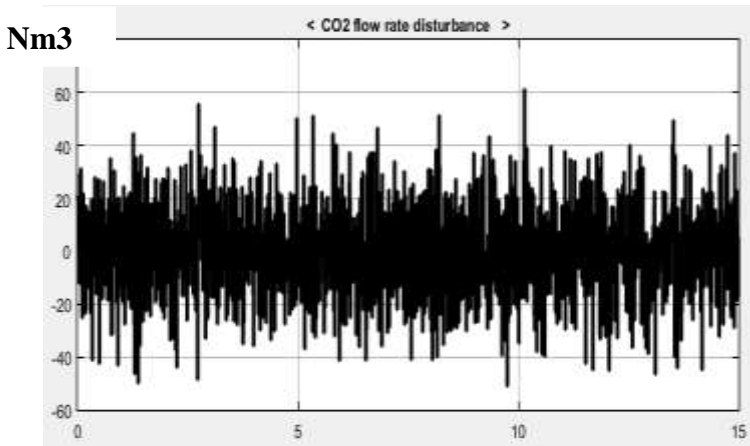


Fig.6. CO2 flow rate disturbance. T(Sec)

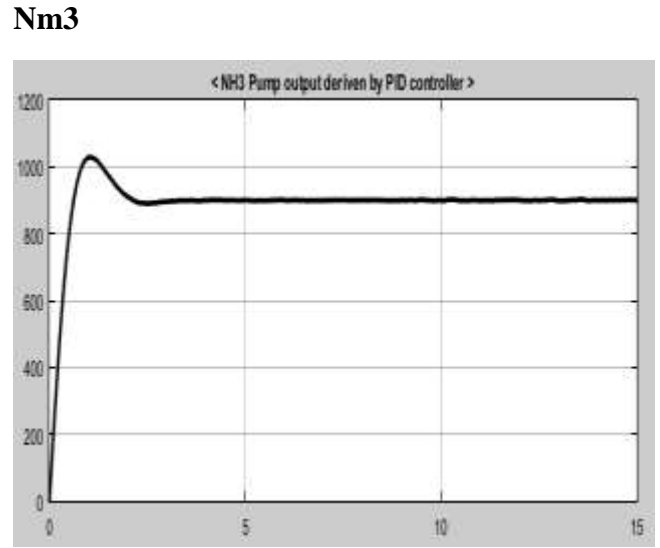


Fig.8. NH3 pump output driven by PID controller. T(Sec)

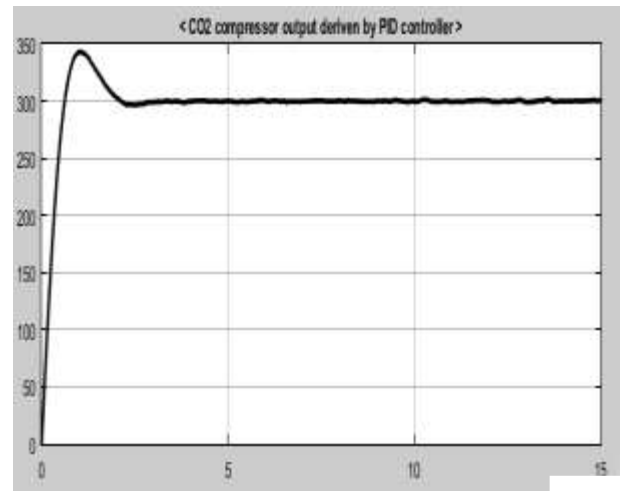


Fig.9. CO2 compressor output driven by PID controller. T(Sec)

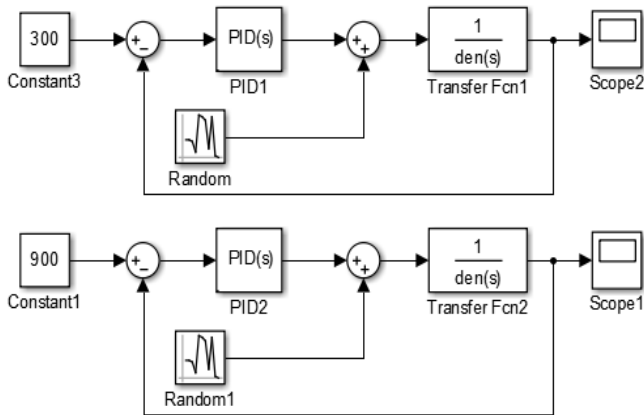


Fig.7. Simulink block diagram of the suggested PID controller.

The Simulink block diagram of the urea reactor model shown in fig.4. The disturbance of NH3 pump is shown in fig.5. Also, for CO2 Compressor is shown in fig.6. For the proposed PID controller shown in fig.7. It has proportional gain K_p as 2.33, the integral gain K_i as 6.67, and the derivative gain K_d as -0.048. The proposed controllers were suggested and tested in simulation using the matlab Simulink toolbox.

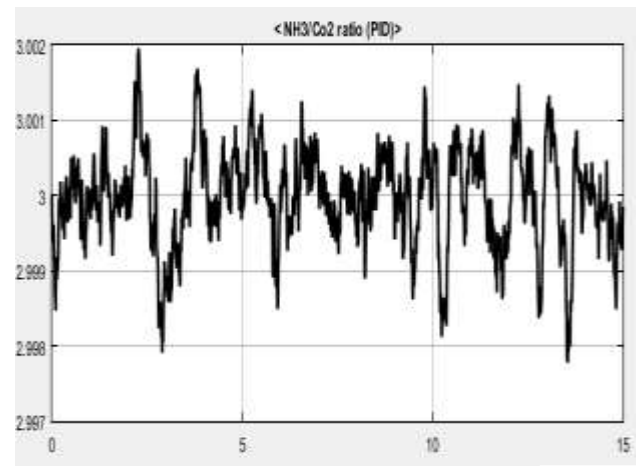


Fig.10. NH3/CO2 ratio curve driven by a PID controller. T(Sec)

Nm3

Dynamic characteristic values. For the proposed PID controller. It has delay time as 0.243s, the rise time as 0.47s, the peak time as 1.02s, the overshoot percentage as 14.3%, and the settling time as 2.9s, NH3 flow rate is shown in fig.8. CO2 flow rate is shown in fig.9. The NH3/CO2 ratio is shown in fig.10.

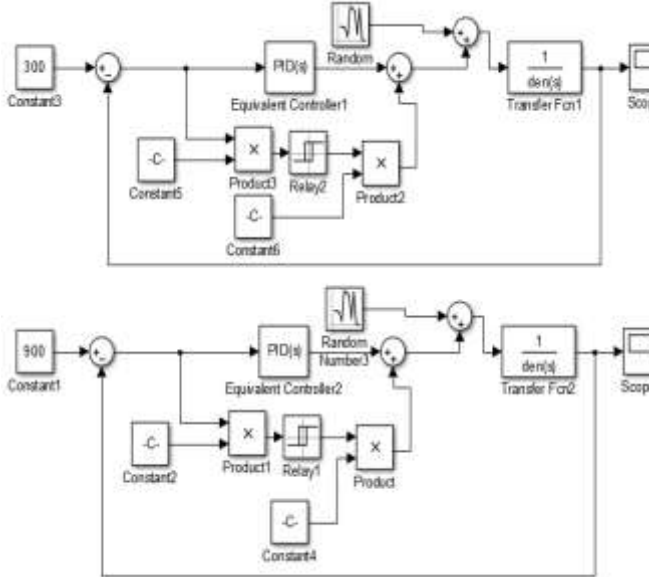


Fig.11. Simulink block diagram of the suggested SMC.

Nm3

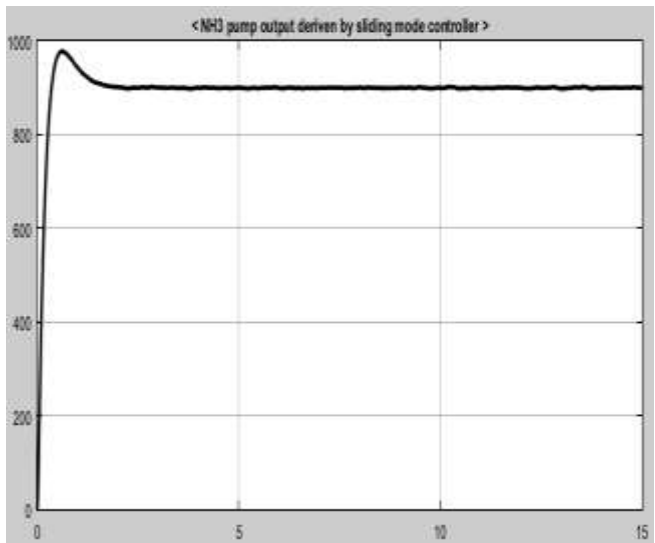


Fig.12. NH3 pump output driven by PID controller. T(Sec)

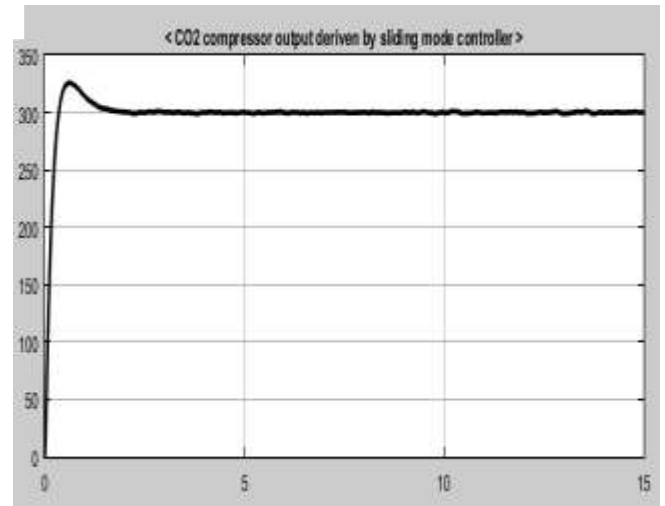


Fig.13. CO2 compressor output driven by SMC. T(Sec)

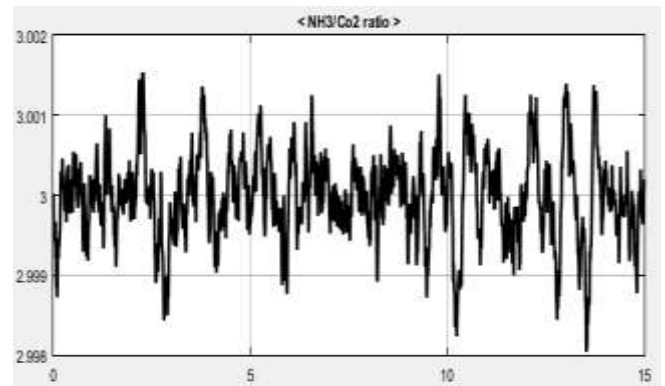


Fig.14. NH3/CO2 ratio curve driven by a SMC. T(Sec)

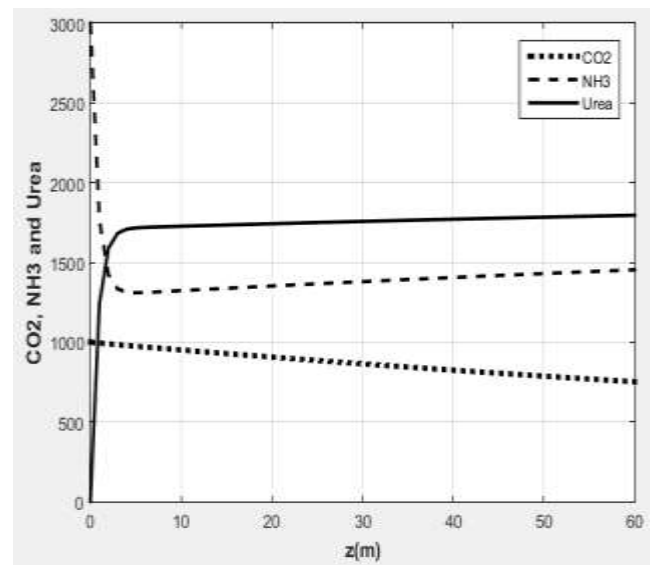


Fig. 15. Urea, NH3 and CO2 Flow Rates

For the sliding mode shown in fig.11. It has delay time as 0.1148s, the rise time as 0.24s, the peak time as 0.6s, the overshoot percentage as 8.6%, and

the settling time as 1.86s, NH₃ flow rate is shown in fig.12. CO₂ flow rate is shown in fig.13. The NH₃/CO₂ ratio is shown in fig.14. The Urea, NH₃ and CO₂ Flow Rates shown in fig.15.

	PID	Sliding mode
Delay time	0.243s	0.1148s
Rise time	0.47s	0.24s
Peak time	1.02s	0.6s
Overshoot percentage	14.3%	8.6%
Settling time	2.9s	1.86s

Table 1. Comparative analysis between dynamic characteristic of PID and SMCs.

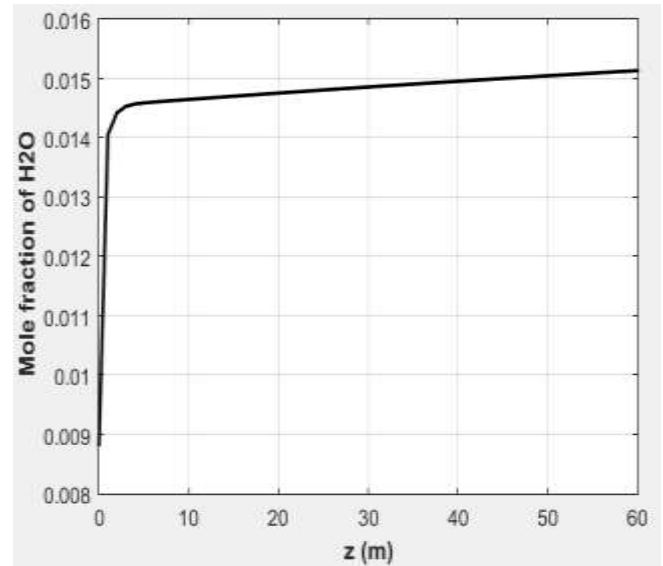


Fig.18. Mole fraction of H₂O.

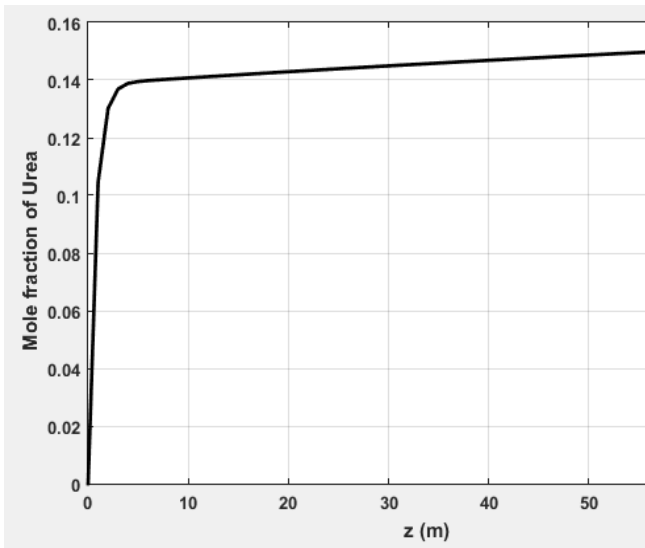


Fig.16. Mole fraction of urea.

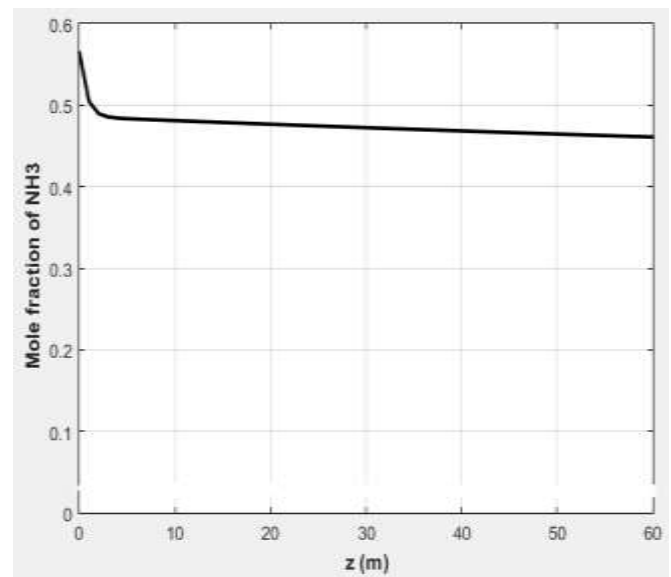


Fig.19. Mole fraction of NH₃.

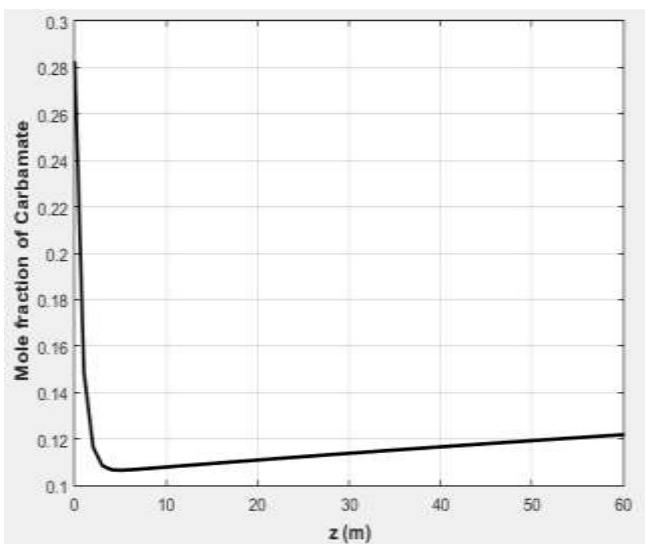


Fig.17. Mole fraction of carbamate.

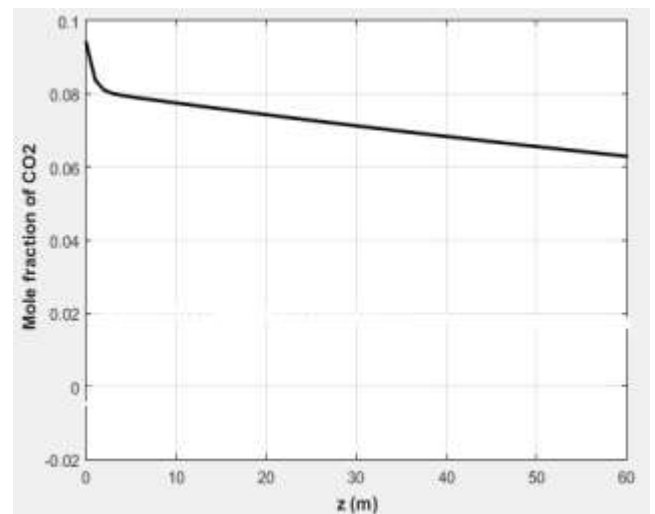


Fig.20. Mole fraction of CO₂

All previous results is shown in table 1. The mole fractions of urea, carbamate, H₂O, NH₃, and Co₂ are shown in Figs (16- 20). Finally, the proposed SMC is better choice to enhance the NH₃/CO₂ ratio as well as the PID proposed. Also, the simulation of applying SMC to maximize the production of urea are represented. The setpoints of the reactor temperature, reactor pressure and NH₃/CO₂ were kept on 182 °C, 145 bar and 3 respectively. When the setpoint which NH₃ flow rate added equal 3.5, the NH₃ flow rate was changed from 77 ton/hr to 80.5 ton/hr. the flow rate of NH₃ changed by SMC is shown in fig.21. The new production of urea was increased from 242 ton/hr to 255.01 ton/hr. The increment of the urea production is shown in fig.22.

We have obtained an increment of 13.01 ton/hr of urea when increased the flow rate of NH₃ from (77 to 80.5 ton/hr). The controlled variables (reactor temperature, reactor pressure, and NH₃/CO₂ ratio) are kept close to their setpoints. The reactor temperature is shown in Fig.23. The reactor pressure is shown in fig.24. Also, the NH₃/CO₂ ratio is shown in fig.25. So the settling time for the reactor temperature is 2.55 hr, for the reactor pressure is 2.56 hr, and for the NH₃/CO₂ is 2.55 hr. The maximum peak for the reactor temperature is 0.202 °C, for the reactor pressure is 0.155 bar, and for the NH₃/CO₂ is 0.004.

Ton/hr

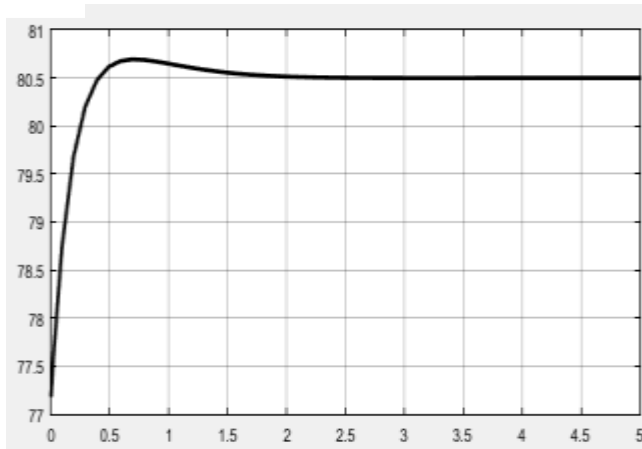


Fig.21 .NH3 flow rate. T(hr)

Ton/hr

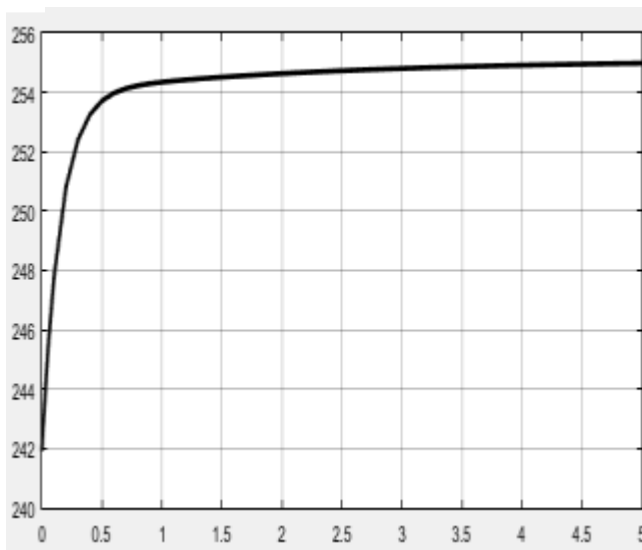


Fig.22 .Urea solution flow rate. T(hr)

C°

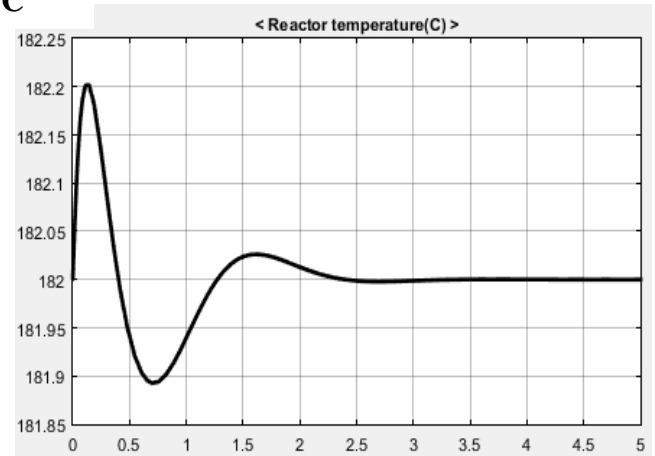


Fig.23 . Reactor temperature.

bar

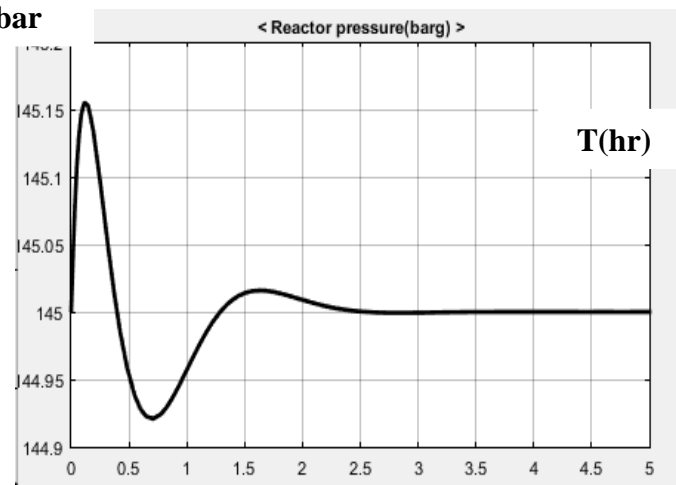


Fig.24 .Reactor pressure. T(hr)

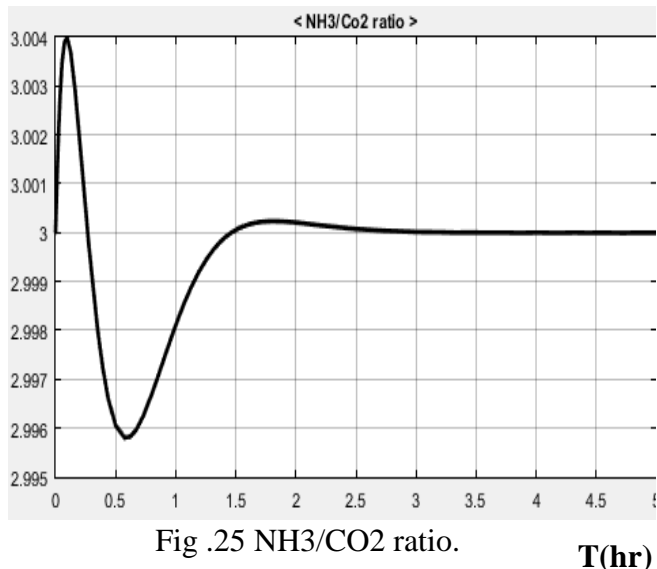


Fig .25 NH3/CO2 ratio.

7. A COMPARTIVE ANALYSIS WITH LITERATURE.

Since SMC is effective that PID a comparison analysis between SMC and literature. For O. M. Agudelo et al. [12], presented a model predictive control (MPC) to maximize the production. It is assumed that increase the NH3 flow rate from 72.2 ton/hr to 80.5 ton/hr. the new production of urea was increased from (241.82 to 253.63) ton/hr. it is obtained an increment of 11.81 ton/hr. in the production of urea. Also the settling time for the reactor temperature is 14 hr, for the reactor pressure is 10 hr, and for the NH3/CO2 is about 14hr, The maximum peak for the reactor temperature is 0.3°C , for the reactor pressure is 0.88 bar , and for the NH3/CO2 is 0.0053.[12].

Table 2. Presents comparison analysis between proposed SMC and O. M. Agudelo et al.

		O. M. Agudelo et al.	P. M. Schuberl, J. S. Higgins,
Urea incremented		11.81 ton/hr	13.00 ton/hr
Maximum peak	Reactor temperature	0.3°C	0.2°C
	Reactor pressure	0.88 bar	0.95 bar
	NH3/CO2 ratio	0.0053	0.004
Settling time	Reactor temperature	about 14hr	2.55 hr
	Reactor pressure	about 11hr	2.56 hr
	NH3/CO2 ratio	about 14hr	2.55 hr

Table2: Comparative analysis with previously related work

From table 2 the proposed SMC is better to increasing the production of urea, reducing the

pollution effect, improving stability, and kept the variable (reactor temperature, reactor pressure, and NH3/CO2 ratio) on setpoints as well as the O. M. Agudelo et al [12].

8. CONCLUSION

The kinetic model for urea synthesis reactor in industrial scale was developed and simulation by matlab Simulink toolbox in the present paper. A modified approach for a SMC is suggested to minimize the NH3 pump and CO2 compressor disturbances in order to reduce the pollution effect in such urea reactor and reduce the unreacted NH3 and CO2, and a PI controller was tested to compare with the sliding mode one. Also, the SMC was used to increasing the production of urea, reducing the pollution effect, and improving stability. This controller kept the reactor temperature under control and also the reactor pressure, and the NH3/CO2 ratio. The results of this controller were great and very good of increment the production of urea. The increment of the production of urea was limited by the maximum feeding flow capacity usually limited by the CO2 compressor.

9. REFERENCES

- [1] H. Uchino, et al, “Toyo Urea Process Advanced Process for Lost and Energy Saving”, New York 1996.
- [2] K. Othmer, “Encyclopedia of Chemical Technology”, Vol. 21, John Wiley & Sons 2007.
- [3] M. Dente, M. Rovaglio, G. Bozzano, A. Sogaro, Chem. Eng. Sci. 1992.
- [4] Basaroff, A. J. Prakt. Chem. 1870, 2, 283.
- [5] R. A. Meyers, et al, “Stamicarbon Dioxide Stripping Urea Process”, New York 1986.
- [6] P. M. Schuberl, J. S. Higgins, S. Higgins, Thermochim Acta 2004.
- [7] I. Mavroudis, Hydrocarb. Process 1971.
- [8] M. Frejaques, Chem. Ind. 1948.
- [9] M. A. Sanyal et al, “Modelling urea processes: A new thermodynamic model and software integration paradigm, Chemical Engineers”, 2003.
- [10] Z. umain, et al, “Kinetic Model for Ammonia and Urea Production Processes”, International Conference on Process Systems Engineering, 2013.
- [11] M. Shyamalagowri, et al, “Model Predictive Control Design for Nonlinear Process Control Reactor Case Study”, IOSR Journal, 2013.

- [12] O. Mauricio, et al, "Control of the Synthesis Section of a Urea Plant by means of an MPC Controller", 16th European Symposium on Computer Aided Process Engineering Published by Elsevier B.V, 2006.
- [13] IM. Fahmy, et al, "Real-Time Control of Industrial Urea Evaporation Process Using Model Predictive Control", Chemical Engineering & Process Technology, 2015.
- [14] S.zendeboudi, et al "ADUAL APPROACH FOR MODELLING AND OPTIMISATION OF INDUSTRIAL UREA REACTOR, THE CANADIAN JOURNAL OF CHEMICAL ENGINEERING, 2013.
- [15] z.Yu and Y. liang, "Design and Realization of Optimization System of Urea Production Process Based on BP Neural Network", "ICCSM", 2010.
- [16] S. skogestad, et al, "simple analytic rules for model reduction and PID controller tuning", journal of process control, vol.13, 2003.
- [17] W. Chang and J. Yan, "Adaptive robust PID controller design based on a sliding mode for uncertain chaotic system", chaos solitions & fractals, vol.26, 2005.
- [18] I. Eker, "second – order sliding mode control with experimental application", ISA Transactions, vol.49, 2010.
- [19] F.G.Areed, et al, "Decoupled Sliding Mode Control for a Multivariable Nonlinear System", International Journal of Computer Applications, 2012.
- [20] M.Lei Tseng and Mi.Chen" Chattering reduction of sliding mode control by low-pass filtering the control signal",ASJC journal,2010.
- [21] L. Zhang et al"Adaptive Tracking Control for Nonlinear Systems with a Class of Input Nonlinearities", ASJC journal,2015.

TRANSMISSION CONTROL PROTOCOL AND CONGESTION CONTROL: A REVIEW OF TCP VARIANTS

Babatunde O. Olasoji^{1*}, Oyenike Mary
Olanrewaju², Isaiah O. Adebayo³

^{1,2,3} *Mathematical Sciences and Information
Technology Department, Federal University
Dutsinma, Katsina State, Nigeria.*

ABSTRACT

Transmission control protocol (TCP) provides a reliable data transfer in all end-to-end data stream services on the internet. There are some mechanisms that TCP has that make it suitable for this purpose. Over the years, there have been modifications in TCP algorithms starting from the basic TCP that has only slow-start and congestion avoidance algorithm to the modifications and additions of new algorithms. Today, TCP comes in various variants which include TCP Tahoe, Reno, new reno, vegas, sack etc. Each of this TCP variant has its peculiarities, merits and demerits. This paper is a review of four TCP variants, they are: TCP tahoe, Reno, new reno and vegas, their congestion avoidance algorithms, and possible future research areas.

Keywords – Transmission control protocol; Congestion Control; TCP Tahoe; TCP Reno; TCP NewReno; TCP Vegas

1. INTRODUCTION

Transmission control protocol (TCP) is an end-to-end connection-oriented, reliable,

process-to-process, stream-oriented protocol. It was designed to create a virtual tube between two TCPs to send data. Reliability means ensuring that data sent from one host (sender) is received by the other host (receiver) and in a situation where the data gets lost in transit, there is a mechanism in place that informs the sender about it and the data is resent (retransmitted). Sequence number is assigned to each byte of data that is transmitted, positive acknowledgement is also needed to be sent back to the sender. The next byte to be sent to the receiver is contained in the acknowledgement to the sender.

In TCP, buffers are used to store data being sent due to the fact that the sender and the receiver may not be writing and reading at the same speed. There is a sending buffer and a receiving buffer at both ends. Because the receiver reads data on a first come first serve basis, if the rate at which the sender writes or sends data is faster than the rate at which the receiver reads or receives data, over time, this will result to data queue which will eventually lead to congestion and data will start dropping off the queue. This is known as drop tail. In order to ensure this does not happen, both the sender and the receiver would have to be running at the same speed or there should be a mechanism in place to ensure there is no congestion by letting the sender know there is a queue building up, so that the sender can reduce the rate at which it is sending packets. TCP has several flow control and congestion control mechanisms designed to control the flow of data and minimize packet loss respectively.

Therefore, in this paper, various variant of TCP algorithms are considered and a head to head comparison is made to further bring into light the differences.

2. CONGESTION

Congestion usually occurs when the load on the network is greater than what the network can handle, mainly because the rate of sending the packets is faster than the rate of receiving the packets by the receiver.

As highlighted in [1], congestion can occur in all areas of communication technology including modern telecommunication, computer networks, wired and wireless network and the internet. With the rate of increase in the use of these communication technologies, especially the internet, it has become imperative to look into the causes as well as solutions to this problem. Possible factors responsible for congestion as outlined in [1] includes: the input rate being equal to or exceeding the capacity of the output rate and also the bookkeeping performances being slow to perform tasks. They also suggested that increasing the resources (bandwidth) and reducing the load in the network would be solutions to these congestion problems.

Various researchers looked at congestion control from different perspectives, [2] reviewed congestion control mechanism in multimedia streaming, [3] did a comparative study on congestion control techniques in high speed network, [4] did a survey on congestion control for packet switched networks.

One of the functions of TCP is to provide congestion control mechanism, the earliest TCP which [5] refers to as old TAHOE (also known as go-back-n model) has two algorithms, and they are: slow start and congestion avoidance. Other TCP variants developed on the first two algorithms by adding fast transmission and fast recovery. All TCP connections start with the slow start phase. When a sender wants to send data, a flow control is imposed by the sender called congestion windows (cwnd), this is based on the sender's assessment of perceived network congestion. The receiver also imposes a flow control called receiver's windows (rwnd), the receiver's window has to do with the amount of available buffer space at the receiver for this connection. The sender considers the congestion window size and the receiver-advertised window. The minimum of these window sizes is considered the actual window size [6]

2.1.SLOW START: Slow- Start algorithm introduced congestion window (cwnd) to the sender's TCP, it is used by the sender as a flow control mechanism [6], in other words it is a sender-based flow control. The slow start phase starts with the sender sending one packet, thereby making the size of the congestion window (cwnd) one packet. The congestion window's size increases exponentially when an acknowledgment of the sent packet(s) has been received by the sender [7]. A threshold is set called Slow-Start threshold (ssthresh), the essence of the threshold is to stop the slow-start phase. The slow-start

phase stops when the size of the window reaches the set threshold and the congestion avoidance phase begins. Slow-start does not prevent congestion but it prevents immediate congestion by making sure a sender does not send a large file [8]

2.2. CONGESTION AVOIDANCE:

When the window size reaches the set threshold, the slow-start phase stops and the congestion avoidance phase begins. Congestion avoidance also known as additive increase prevents congestion from happening fast by slowing down the rate of increase of congestion window from exponential increase in slow-start algorithm to linear increase in congestion avoidance algorithm. The result of this is that the rate of growth of the congestion window is reduced as it increases linearly by one segment every round trip time. When a packet is sent, a timer (retransmission timer) is set, if an acknowledgement is not received by the sender from the receiver before the expiration of the timer, the sender will assume there is congestion and the packet has been lost, thereby making the sender to retransmit the packet. The communication path (pipeline) is also emptied and the slow-start algorithm process restarts again i.e the cwnd is set to one segment.

2.3. FAST TRANSMISSION AND

FAST RECOVERY: There are two ways TCP detects congestion, they are: Retransmission time-out and three duplicate acknowledgements [9]. Congestion detection through time-out, empties the communication path (pipeline) by reducing the cwnd to one segment thereby causing a major cost in high band-width delay product links, due to this shortcoming, congestion avoidance algorithm was modified to avoid this. [6, 9 & 10] explained how three duplicate acknowledgement works and how it detects congestion before retransmission timer expires. Fast retransmission is a modification or an enhancement of congestion avoidance algorithm where congestion is detected through three duplicate acknowledgements. In fast retransmission, slow-start is not performed after the missing packet has been resent, instead, congestion avoidance is performed, with this, the communication line is not emptied and the whole process of slow-start is not restarted, this is known as the fast recovery algorithm.

We shall start the paper by taking a brief look at each of the congestion avoidance algorithms and noting how they differ from each other. In the end we shall do a head to head comparison to further bring into light the differences.

3. TCP VARIANTS

3.1. NEW TAHOE

Early variant of TCP implemented two congestion control algorithms which are: Slow-Start and congestion avoidance algorithm, this variant is referred to as the OLD TAHOE [5], in old tahoe, congestion is assumed when the sender has not received the acknowledgement of the sent packet from the receiver before the retransmission timer expires (timeout). Due to the shortcomings of congestion detection through timeout, TCP TAHOE was designed, it adds a new algorithm to the existing congestion control algorithms (i.e. Slow-Start, congestion avoidance) called fast retransmit [5]. In tcp tahoe, congestion is assumed when three duplicate acknowledgements (Dupack) are received by the sender, the lost segment is retransmitted without waiting for the expiration of the retransmission timer timeout, it thereafter resets the congestion window to one packet and begins slow-start. This causes higher utilization and connection throughput [11]

3.2. TCP RENO

TCP reno replaces the slow start with congestion avoidance by reducing the congestion window to one half after a lost packet has been retransmitted using fast retransmit algorithm, this prevents the communication path from becoming empty. This is achieved by adding a new algorithm

to tcp tahoe known as fast recovery. [12, 13] analyzed the mechanisms of fast recovery algorithm explicitly, according to [13], tcp reno solves many of tcp tahoe's problem. The problem with tcp reno is that when multiple packets are dropped, it will have performance problem, which means tcp reno performs well when only one packet is lost as it can only retransmit at most one packet.

3.3. TCP NEW RENO

Due to the shortcoming of TCP reno, a slight modification in fast recovery algorithm became necessary [13]. TCP reno works well when only one packet is lost but when multiple packets are lost then it will have performance problem, because when a partial acknowledgement is received by the sender, it brings the sender out of fast recovery which result in a timeout [10]. The result of the modification is called TCP new reno. The fast recovery algorithm was modified such that when a sender receives a partial ACK, it does not come out of fast recovery, according to [13], the sender could deduce whether multiple losses have occurred in the same window from the duplicate acknowledgement (dupack) received. Thus, TCP sender continues to retransmit lost segments when it receives a partial ACK and when it receives a full ACK, it sets the congestion window to the threshold value and invokes the congestion avoidance algorithm.[9]

3.4. TCP VEGAS

TCP VEGAS is a modification of TCP RENO, it capitalizes on the shortcoming of

this TCP variant, which includes its congestion detection and control mechanism that uses the loss of segment as a signal that there is congestion in the network. These modifications have made TCP VEGAS to achieve between 40 and 70% better throughput [14]. The modifications made are:

1. A modified slow-start mechanism.
2. An improved congestion avoidance mechanism.
3. The use of a new retransmission mechanism.

TCP Reno's congestion detection and control mechanism uses the loss of segment as a signal that there is congestion in the network, it has no mechanism in place to detect incipient stages of congestion to prevent congestion, rather it reacts to it after it has happened, hence TCP Reno is a reactive TCP variant. On the other hand, TCP Vegas uses a more sophisticated bandwidth estimation scheme which attempts to avoid congestion rather than react to it after it has happened. To achieve this, TCP Vegas uses a more accurate RTT calculation, it reads and records the system clock each time a segment is sent, such that when an ACK is received, Vegas reads the clock again and does the RTT calculation using the time and the timestamp recorded for the relevant segment. Using an accurate RTT estimate leads to a more accurate calculation of the number of data packets that a source can send [14].

TCP Vegas measures and controls the amount of extra data Δ the connection has in transit, the idea is that the extra data would not have been sent if the bandwidth used by the connection exactly matched the

available bandwidth in the network. Vegas maintains the right amount of extra data in the network.

The extra data is the difference between the expected throughput rate and the actual throughput rate of the sent data.

The formula for the Expected throughput is given by;

$$Expected = CWND / BaseRTT \quad [15]$$

The *CWND* is the current congestion window size, while the *BaseRTT* is the minimum of all measured round Trip Time. Usually, the RTT of the first segment sent by the connection.

The formula for the Actual throughput is given by:

$$Actual = CWND / RTT \quad [15]$$

Where RTT is the round trip time of the current segment.

TCP Vegas compares the actual throughput to the expected throughput and it adjusts the CWND accordingly.

TCP VEGAS does not use the traditional slow-start mechanism because it induces packet losses to estimate the available bandwidth in the network, instead, it uses the difference (also known as extra data Δ) between the expected and the actual throughput rates to estimate the available bandwidth in the network. It increases the window size exponentially only every other return trip time (RTT). In between, the CWND remains fixed so that a valid comparison of the Expected and Actual sending rates can be made.

$$\Delta = (Expected - Actual) * BaseRTT \quad [15]$$

A threshold γ is set for the difference (Δ) between the expected rate and the actual rate such that when the difference goes above the threshold γ , it, reduces its CWND by one-eighth, exits slow-start and enters the congestion avoidance phase[15].

During congestion avoidance phase in TCP reno, congestion window increases linearly after each RTT which eventually will lead to congestion, TCP VEGAS on the other hand tries to avoid the occurrence of congestion by not increasing continually the congestion window[15], it also detects incipient congestion by comparing the actual rate with the expected rate of the throughput and adjust the CWND accordingly. Two thresholds α and β are set as constant, this is done to ensure that the extra data is kept in check. The two thresholds mean having too low or too much extra data in the network, respectively. If the difference (extra data Δ) is greater than β , it is considered to indicate incipient congestion and the CWND is reduced linearly during the next RTT. If Δ is lesser than α , the CWND is increased linearly during the next RTT. If the Δ lies between the two thresholds (α and β), the CWND remains constant.

$$CWND = \begin{cases} CWND + 1, & \text{if } \Delta < \alpha \\ CWND - 1, & \text{if } \Delta > \beta \\ CWND, & \text{if } \alpha \leq \Delta \leq \beta \end{cases} .$$

[15]

Since TCP VEGAS does not use packet loss to detect congestion, it uses decrease in sending rate as compared to the expected rate. When the actual rate is too far from the expected rate, it increases transmissions to make use of the available bandwidth, but when the actual rate comes too close to the expected value, it decreases its transmission

to avoid over saturating the bandwidth, the difference in the bandwidth estimation schemes enables TCP VEGAS to utilize the available bandwidth more efficiently.

Basically, TCP reno uses two mechanisms to detect and retransmit lost segments. The first one is the retransmission timeout which is integrated in the original mechanism of the TCP specification. It uses a coarse-grain timer. Hence, it retransmits when a coarse grain timeout occurs. According to [14], the coarse grained timer estimation of RTT is not very accurate and it also causes an unnecessary delay. The second mechanism is when the sender receives and duplicate ACKs (n is usually 3). TCP Vegas uses fine grained RTT measurement, where a timeout period is computed for each packet. This ensures an accurate RTT measurement [15] TCP vegas achieves between 40 and 70% better throughput and it outperforms other implementations of TCP in many cases. TCP vegas is able to achieve an improved throughput not by an aggressive retransmission strategy where bandwidth are stolen away from TCP connections but by the effective use of the available bandwidth. However, in an heterogenous network where TCP vegas connection has to compete with other connections that use TCP reno, it does not receive a fair share of bandwidth due to its proactive congestion avoidance mechanism and the aggressive nature of TCP reno. Also, even though it has been available for a few years, it has not been widely adopted due to its perceived incompatibility with TCP reno [15].

4. CONCLUSION

TCP Tahoe, Reno and New Reno all use coarse grained timer to calculate their RTT, the RTT calculated is usually not accurate and causes a lot delay. They also detect congestion through packet loss (Loss-based Congestion Avoidance (LCA)), they have no mechanism in place to detect incipient stages of congestion to prevent congestion, rather they react to it after it has happened which make them reactive TCP variants. It became important to calculate RTT using fine grained timer to calculate the RTT in order to avoid delay and to also detect congestion without any loss of packets. A modification to TCP Reno was made and the result of this modification was named TCP Vegas. TCP Vegas implements fine-grained timer in RTT estimation of the retransmission mechanism. It gives accurate RTT and it doesn't detect congestion through packet loss (Delay-based Congestion Avoidance (DCA)) this makes it achieve between 40 and 70% better throughput and it outperforms other implementation of TCP Variants. However, it has not been widely implemented for many reasons which include its incompatibility with TCP Reno. Also, TCP Vegas looks fine in a homogenous network but in a heterogeneous network where other TCP variants are implemented, there would be unfairness in the sharing of bandwidth due to their aggressive nature [16].

Due to the shortcomings of TCP vegas, further work will be carried out on it in subsequent research work.

REFERENCES

1. Subramani B., Karthikeyan T.,(2014). "A Review on Congestion Control." International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 1, January 2014.
2. Rajarajeswari S., Sutha J., (2013). "A survey on multimedia streaming congestion control mechanism." International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 2, February 2013)
3. Shakeel A., Adli M., et al (2009). "Comparative study of congestion control techniques in high speed network."IJCSIS) International Journal of Computer Science and Information Security, Vol.6, No. 2, 2009
4. Socrates C., Devamalar P.M., Sridharan R. K., (2014). "Congestion Control for Packet Switched Networks: A Survey." International Journal of Scientific and Research Publications, Volume 4, Issue 12, December 2014 1 ISSN 2250-3153
5. Rohan S., Suman S., Viswanathan P. (2014), "Transmission Control Protocol: Comparison of TCP Congestion Control Algorithms using NetsimTM "

6. Stevens W., (1997) "TCP slow start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms." RFC 2001.
7. Jacobson, V. Congestion avoidance and control. In *Proceedings of SIGCOMM '88* (Stanford, CA, Aug. 1988), ACM.
8. Socrates C., Devamalar P.M., Sridharan R.K., (2014), "Congestion Control for Packet Switched Networks: A Survey". International Journal of Scientific and Research Publications, Volume 4, Issue 12, December 2014 1 ISSN 2250-3153.
9. Torkey H., Attiya G., Morsi I. Z., (2012), "Modified Fast Recovery Algorithm for Performance Enhancement of TCP-NewReno." International Journal of Computer Applications (0975 – 8887) Volume 40– No.12, February 2012.
10. Fahmy S. and Karwa T,(2001), "TCP Congestion Control: Overview and Survey Of Ongoing Research." *Computer Science Technical Reports*. Paper 1513. <http://docs.lib.purdue.edu/cstech/1513>
11. Shamimul Q., Kumar M., (2010), "Impact of Random Loss on TCP Performance in Mobile Ad-hoc Networks (IEEE 802.11): A Simulation-Based Analysis."(IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 1, 2010.
12. K. Fall, and S. Floyd, "Simulation-Based Comparison of Tahoe, Reno and SACK TCP." *Computer Communications Review* ACM SIGCOMM, Vol. 26, No. 3, July 1996
13. S. Floyd and T. Henderson. The NewReno modification to TCP's fast recovery algorithm. RFC 2582, April 1999. Also see http://www.aciri.org/flfloyd/tcp_small.html.
14. Lawrence S. Brakmo, Sean W. O'Malley, and Larry L. Peterson TCP Vegas: New Techniques for Congestion Detection and Avoidance, SIGCOMM' 94, 1994.
15. Chan Y., Lin C., Chan C., Ho C., (2010) "CODE TCP: A competitive delay-based TCP."
16. Yew B., Ong B., Ahmad R., (2011). "Performance Evaluation of TCP Vegas versus Different TCP Variants in Homogeneous and Heterogeneous Networks by Using

Network Simulator 2.” International
Journal of Electrical & Computer
Sciences IJECS-IJENS Vol: 11 No:
03.

DETECTION OF BLACK HOLE ATTACKS IN MANETS BY USING PROXIMITY SET METHOD

K. Vijaya kumar¹, Dr. K.Somasundaram²,

1. Research Scholar (Karpagam University), Assistant Professor, Department of Computer Science Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India.

2. Professor, Department of Computer Science and Engg., Vel Tech High Tech Dr.RR Dr.SR Engineering College, Avadi, Chennai, Tamilnadu India.

ABSTRACT

A Mobile Adhoc Networks (MANETS) is an infrastructure less or self configuring network which contain a collection of mobile nodes moving randomly by changing their topology with limited resources. These Networks are prone to different types of attacks due to lack of central monitoring facility. The main aim is to inspect the effect of black hole attack on the network layer of MANET. A black hole attack is a network layer attack also called sequence number attack which utilizes the destination sequence number to claim that it has a shortest route to reach the destination and consumes all the packets forwarded by the source. To diminish the effects of such attack, we have proposed a detection technique by using Proximity Set Method (PSM) that efficiently detects the malicious nodes in the network. The severity of attack depends on the position of the malicious node that is near, midway or far from the source. The various network scenarios of MANETS with AODV routing protocol are simulated using NS2 simulator to analyze the performance with and without the black hole attack. The performance parameters like PDR, delay, throughput, packet drop and energy consumption are measured. The overall throughput and PDR increases with the number of flows but reduces with the attack. With the increase in the black hole attackers, the PDR and throughput reduces and close to zero as the number of black hole nodes are maximum. The packet drop also increases with the attack. The overall delay factor varies based on the position of the attackers. As the mobility varies the delay and packet drop increases but PDR and throughput decreases as the nodes moves randomly in all directions. Finally the simulation results gives a very good comparison of performance of MANETS with original AODV, with black hole attack and applying proximity set method for presence of black hole nodes different network scenarios.

KEYWORDS:

AODV protocol, security, black hole attack, NS2 simulator, proximity set method, performance parameters.

1. INTRODUCTION:

A MANET [3, 4, 5] is a self configuring network which contains a collection of nodes which deploys multiple hop packet radio service to communicate among themselves without any infrastructure or centralized control. The various applications of MANETs are emergency, military, battle field etc in which they could share surveillance data in order to improve the precision and efficiency of attack [5] and chance of survival. Due to the basic characteristics of MANET 's like node instability, limited computation resources, bandwidth, and power supply, constant topology change, distributed operations, and lack of centralized management, it is demanding for the design and implementation of network applications and protocols for a mobile ad-hoc network. These networks are vulnerable to several types of attacks namely active and passive attacks. Active attacks are those attacks which try to interrupt the proper functionality of the network. Passive attacks are very dangerous when compared to active attacks because it will not change or alter the normal functionality of network whereas the attackers try to listen silently or retrieve the important information available inside the data packets. The security for routing protocols [11] of MANET's can be obtained mainly in two major ways like Prevention as well as detection and reaction. The instable of ad hoc networks requires that prevention techniques should be balanced by detection techniques, which observe the security status of the network and identify malicious behavior of the nodes [6]. The malicious behavior of a node cannot be identified easily when a set of nodes in MANET compromised which is a grave

problem. The intention of these nodes is to generate new routing messages to advertise non-existent links, to provide incorrect link state information, and flood other nodes with routing traffic, thus inflicting failure in the network[6]. The most widely used routing protocols in MANETs is the ad hoc on-demand distance vector (AODV) routing protocol which works on source initiated on-demand routing protocol. AODV is much exposed to the well known black hole attack [2].

2. RELATED WORK

In [1] the author Sumathi specified Energy based secure protocol E-AODV which provides data transfer using digital signature algorithm by enhancing the throughput and security and even though the node increases because of its Trust and Energy model used he proposes security enhancement. It is assured that using back propagation algorithm with gradient descent based learning ensures CIA triangle in the proposed E-AODV protocol. An ANN approach is used to check the effectiveness of the proposed system and is analyzed in terms of mean square error value, learning rate, gradient value and finally the back propagation network also ensures CIA Triangle security in E-AODV.

Christeena Joseph [3] said that when the attacker is near the source the impact is severe than it is farther. Similarly as the number of black hole increases, PDR and throughput decreases. This is due of the fact that the black hole sends the RREP with highest destination sequence number without verifying for a route in its routing table. The Intrusion Detection Systems (IDS) can detect whether the network is under an attack, notify the network and hence able to isolate the attacker. The anomaly detection system has advantage over signature and specification IDS is that it can detect unknown attacks.

Senthil Murugan, the author [4] described an Intrusion Detection System (IDS) implemented using Genetic Algorithm and tested with networks of varied node configurations. The algorithm will be tested for more number of nodes and the performance analysis will be done in terms of execution time and efficiency of the algorithm as the node number is increased. This can be extended

to DSDV protocol for detecting black hole to avoid routing through the attacker.

The author [5] proposed a method by way discovery which is focused around AODV utilizing recreations created within Network Simulator to shield again the worm gap attacks in remote unprepared systems and here wormhole attack is caught without any stuff, area data and clock synchronization. At last it enhanced Packet Delivery Ratio, Average postponement, Packet-misfortune, Detection-degree think about than other wormhole attacks.

Bhakte and Dr Rahul [6] said that security issues have been ignored while designing routing protocols for ad-hoc networks. AODV protocol is susceptible to many malicious attacks including Black Hole Attacks. The proposed protocol said Secure Route Discovery and Data Transmission from Black Hole Attacks on AODV-based Mobile Ad-hoc Networks is the mechanism that uses the cryptographic technique for securing route discovery and data transmission thus the packet loss will be reduced.

In the Proposed solution the author Rahul Vasant Chavan [7] tried to convert Black hole node in to normal node, which selectively Perform black hole attacks by deploying IDSs System in MANETs. By evaluating the amount of abnormal difference between RREQs and RREPs transmitted from the node, all IDS nodes going to define malicious behavior in MANET, which estimates the Packet Threshold Value of a node which results that the percentage of data packet loss is better than DSR in presence of multiple Black hole nodes.

Gomathi .et.al [8] specifies the uniqueness in MANET make more susceptible than wired network. In this the author tried to categorize MANET attacks based on various characteristics and layer based attacks are analyzed. However different security mechanisms are employed to prevent these attacks ranging from Intrusion Detection System (IDS) to various cryptographic algorithms.

C.V. Anchugam and K. Thangadurai [9] analyzed the effects of black hole attack in the light of network load, throughput and end-to-end delay in MANETs and simulating the black hole attack

using reactive routing protocols and said that AODV without attack gives better result in all situations. The author specifies that under attack case system has more packet drop ratio it is always greater to threshold. The author designed and implements a security algorithm for detection of black hole attack based on Ad hoc On-Demand Distance Vector routing protocol and Ant Colony Algorithm.

The author [10] concluded that the probability based algorithm provides a significant approach for constructing black hole free network without any overhead for using a separate procedure to detect a black hole node and then for preventing it. A significant improvement in all the parameters can be observed, PDR, E2E Delay or Throughput. This algorithm can be used as in other approaches towards black hole detection or prevention and also can be used to study of cooperative black hole attack.

Nakka Nandini [12], the author described the redundant route method, we send the ping packet and find the existence of two routes, and find the best route among them. In another method, i.e. detection, prevention and reaction AODV (DPRAODV) by updating the threshold values and changing them between the sequence number and the RREP packets, we can prevent the black hole attack to some extent in the networking environment.

The author [13] proposed the Consensus based algorithm helps to detect and prevent the malicious nodes and also provides the successful transmission of the packets between the nodes. This increases the network's performance, routing and the throughput and the results are carried out using ns2.

In this paper Kaur [15] propose a scheme for detecting black hole attack in MANETs namely clustering based DSR Protocol which is introducing clustering in the route discovery phase of DSR protocol. The proposed protocol is simple and efficient and also provides better values for packet drop ratio and detection rate as compared to existing scheme in simulation results.

Rashmi, Ameeta Seehra [18] specified a light weight solution which is based on simple acknowledgement scheme to prevent black-hole

nodes in MANET. In this the mobile check-points detect the presence of malicious nodes in the source route and with the help of intrusion detection system the suspected nodes are obscured from the network. The simulation results show that clustering approach is responsible for full delivery of packets even in presence of multiple black-hole nodes.

In the paper, the author [25] projected the black hole detection and correction algorithm for MANETs (OBHDPA) which improves the performance of the network structure by 50 percent. This result has been verified using the ns-2 simulator. The proposed solution can be applied to identify multiple black hole nodes, discover the secure paths from source to destination by preventing black hole problem.

In [26] the author said that a black hole attack caused by a malicious node impersonates a destination node by sending forged RREP to a source node that initiates route discovery, and consequently deprives data traffic from the source node. The author presented a new detection method based on dynamically updated training data and analyzed the blackhole attack and introduced the feature in order to define the normal state of the network.

3. EXISTING - AODV ALGORITHM

Adhoc –on-demand distance vector (AODV) routing protocol uses an on-demand approach for finding routing routes, that is route is established only when it is required by a source node or transmitting data packets. It employs destination sequence number to identify the most recent path. In AODV the sender node and the intermediate nodes store the next hop information corresponding to each flow for packet data transmission [21]. In AODV the sender floods *RouteRequest* packet in the network where a route is not available for the desired receiver [3]. It may obtain multiple routes to different receivers from a single *RouteRequest*. In AODV it uses a receiver sequence number (RecSeqNum) to determine an up-to-date path to the receiver. A node updates its path information only if the RecSeqNum of the current packet received is greater than the last RecSeqNum stored at the last node[3].

A *RouteRequest* carries the sender identifier (SenID), the receiver identifier (RecID), the sender

sequence number (SenSeqNum) , the receiver sequence number (RecSeqNum) , the broadcast identifier (BcastID), and the time to live (TTL).RecSeqNum indicates the freshness of the route that is accepted by the sender. When an intermediate node receives a *RouteRequest*, it either forwards it or prepares a *RouteReply* if it has a valid route to the receiver[16]. The validity of a route at the intermediate is determined by comparing the sequence number at the intermediate node with the receiver sequence number in the *RouteRequest* packet. If a *RouteRequest* is received multiple times, which is indicated by the BcastID-SenID pair, the duplicate copies are discarded. All intermediate nodes having valid routes to the receiver, or the receiver node itself, are allowed to send *RouteReply* packets to the source. Every Intermediate node, while forwarding a *RouteRequest* , enters the previous node address and its BcastID. A timer is used to delete this entry in case a *RouteReply* is not received before the timer expires. This helps in storing an active path at the intermediate node as AODV does not employ sender routing of data packets[9]. When a node receives a *RouteReply* packet, information about the previous node from which the packet was received is also stored in order to forward the data packet to this next node as the next hop toward the destination. When a link break is detected in any active route RERR message is generated. Some essential fields of RouteRequest, RouteReply and RouteError [28] messages are

Sender's IP address
remote Receiver's IP Address
remote Receiver's Sequence Number
Additional remote Receiver's IP Address
Additional remote Receiver's Sequence Number
Receivers Count

Table3 RouteError

4. PROBLEM DESCRIPTION

In MANETS's the various routing protocols faces many security aspects of problems due to the dynamic nature and resource constraints. There are different types of attacks that can occur when the malicious node present in the network is intended to attack directly the data traffic and intentionally drops, delay or alter the data traffic passing through it. One particular attack at network layer is Black Hole Attack which is very dangerous and active attacks on the MANETs[21,23]. It is formed during the week routing infrastructure, when a malicious node joins the network this problem arises. In detection system for ad hoc networks are extremely difficult due to lack of central controller, bandwidth limitations, and dynamic topology in mobile ad hoc networks. A Black Hole Attack [22] is performed by a single node or combination of nodes, also called selfish node. The method how malicious node fits in the data routes varies. The below Figure shows how black hole problem arises, here node "N1" i.e sender nodewants to send data packets to node "N6" receiver and initiate the route discovery process. So if node "N4" is a malicious node then it will claim that it has active route to the specified receiver as soon as it receives RREQ packets. It will then send the response to node "N1" before any other node. In this way node "N1" will think that this is the active route and thus active route discovery is complete. Node "N1" will ignore all other replies and will start sending data packets to node "N4". In this way the entire data packet will be lost consumed or lost. As an outcome, the sender and the receiver nodes became inefficient to communicate with each other. While AODV treats RREP messages having higher sequence number to be fresher, the malicious node all the time send the RREP having higher sequence number. So RREP message, once received by sender node is treated as new, too. The outcome is that there is a high probability of a malicious node effort to organize the black hole attack in AODV. Black hole attack

Sender's IP address
Senders Sequence Number
Receiver's IP Address
Receiver's Sequence Number
RouteRequest Id number
Originator IP Address
Originator Sequence Number
No. of Hops

Table 1 RouteRequest

Sender's IP address
Receiver's IP Address
Receiver's Sequence Number
Originator IP Address
No. of Hops
TTL

Table2 RouteReply

problem in MANETs could be very serious security problem to be resolved.

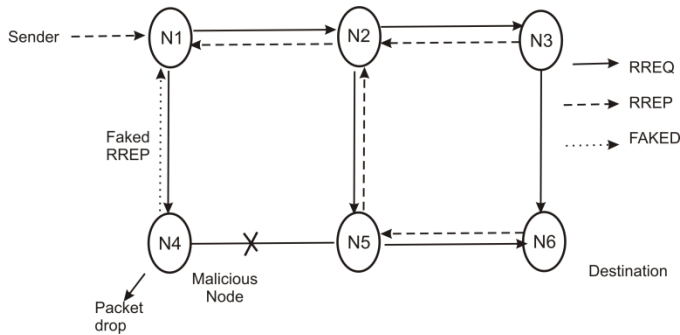


Figure 1 Function of malicious node

5. PROXIMITY SET METHOD (PSM)- PROPOSED METHODOLOGY

The main objective of this paper is to keep focusing on detecting black hole attack which can be stated of two types. The former type of black hole attack can occur when the malicious node on the pathway directly attacks the data transfer by purposely discarding, delaying or varying the data traffic passing through it. This type of black hole attack can be easily mitigated by setting the promiscuous form of each node and listening to see if the next node on the path forward the data traffic as expected. The other type of black hole attack is the attack is done on routing control traffic. The malicious node can act as if to be like some other node and advertise it having the shortest path to the source node whose packets it is interested in. In this way, this malicious node becomes a black hole since the data traffic is misrouted to an inaccurate route or destination. We developed methods to detect this type of routing misbehavior caused by black hole attack.

We propose a proximity set-based method (PSM). PSM can be briefly elaborated as, once the normal path discovery procedure in a routing protocol is completed, the source node sends a special control packet to request the destination to send its current proximity set. By comparing the received proximity sets, the source node can determine whether there is a black hole attack in the network. To diminish the impact of the black hole attack, we design a routing recovery protocol which identifies the black hole or malicious nodes to establish the path to the correct destination.

We use reactive AODV protocol as the routing protocol in implementing proximity set method.

Proximity set is defined as all of the nodes that are within the radio communication range of a node. Due to the dynamic rapid moment of the nodes, the proximity set of a node keeps changing and it is expected that the proximity set changes faster when mobility increases. The chance that two mobile nodes have the same proximity set at the same time is very small. So the proximity set provides a good “identity” of a node, i.e., if the two proximity sets received at the same time are different enough, we can conclude that they are generated by two different nodes.

Two processes are implemented to say that determining proximity set of a node is a good identification for finding malicious node.

- In the first process, we calculated the proximity set difference of one node at different time instants t and $t + 1$ under different moving speeds and network density (i.e., number of nodes in the system),
- In the second process, we calculated the proximity set difference of two different nodes, say node A and node B, at the same time. We measured the number of Nodes in the set by using principle of exclusion and inclusion set definition as $((\{A's\ proximity\ set\} \cup \{B's\ proximity\ set\}) - (\{A's\ proximity\ set\} \cap \{B's\ proximity\ set\}))$.

Based on this proximity set information, we design a method to deal with the black hole attack, which consists of two parts: detection and response.

5.1 DETECTION:

In order to collect proximity set information, we introduce two types of control packets in the detection phase:

requestproximityset(RQPS) and
replyproximityset(RPPS).

The packet format of RQPS is as follows:

```
{srcaddr, destaddr, requestproximityseq#, nexthop
}
```

srcaddr is the IP address of the source node S

destaddr is the IP address of the destination D.

Each node is responsible for maintaining one counter: the sequence number of the RQPS, Each time a node sends a RQPS, requestproximityseq# increases by one. The sequence number in each node uniquely identifies the RQPS, which unicast to

the destination using the underlying AODV routing protocol.

D or D' (malicious node), after receiving RQPS, replies a message RPPS.

The message format of RPPS is as follows:
{*srcaddr, destaddr, requestproximityseq#, proximity set*}

The first three items, i.e., *srcaddr, destaddr, requestproximityseq#*, identify to which RQPS this RPPS corresponds.

Proximity set contains the current proximity set of D or D'. This RPPS unicast back to S.

There are two major steps

Step 1: Collect proximity set information.

By using AODV protocol, the source node S floods RREQ packets across the network to find a route to the destination node D. Now for each received RREP, S will unicast a RQPS packet, and the RQPS packet will go to either D or D', depending on the path contained in RREP.

After D or D' receives RQPS, it will generate a RPPS packet, which contains its current proximity set, and unicast it back to S.

Step 2 Determine whether there exists a black hole attack.

The source node S, after receiving more than one RPPS packet in a certain period will start comparing the received proximity sets. The difference among the proximity sets is defined as the union of the received proximity sets minus the intersection of the proximity sets. If the difference is larger than the predefined threshold value, S will know that the current network has black hole attacks and take some actions to respond to it. One concern is that what if D' first requests the proximity set of D, and replies it to S? We think that it is difficult for D' to do so. Because D' claim D's address, D' has to use D's address to request D's proximity set, (otherwise, D's proximity set can find that D' is a malicious). But D will raise an alert to this request, because it uses the same address of D.

5.2 RESPONSE:

We assume there exists a public key infrastructure, which S can use to authenticate D or D'. After S detects the black hole attack, it will use the cryptography-based method to authenticate D and

D'. In this way, S can identify D, the true destination.

Once D is identified, S will send a *modifyrouteentry* MRE control packet to D to form a correct path by modifying the routing entries of the intermediate nodes from S to D. We call this routing recovery protocol. The packet format of MRE is as follows:

{*destaddr, correctpath* }
destaddr is the IP address of D. *correctpath* is the hop by hop path from S to D.

S can get the information *correctpath* from the received RPPS's. After each node receives the MRE, it will modify its corresponding routing entry (identified by the IP address of D) to make its next hop on the path to D, instead of D'. After D receives MRE, a correct path has formed between S and D, which will make the traffic of S go to the correct destination.

6. IMPLEMENTATION AND RESULTS:

The proximity set method is implemented on AODV protocol and simulations are presented on various parameters like average end to end delay, throughput, packet delivery ratio, energy consumption analysis [1,6,10,12] etc and compared on original AODV without any blackhole attack, AODV with blackhole attack and AODV with PSM method.

TABLE 4 SIMULATION PARAMETERS

PARAMETER	VALUE
Simulator	NS2
Topology area	3000 X3000
Number of mobile nodes	49
Number of blackhole nodes	03
Traffic type	CBR-UDP
Movement model	Random way point
Simulation time	300 seconds
Range of transmission	100 m
Payload of data	512 byte
Routing protocol	AODV
Pause time	2 seconds

PDR (Packet delivery ratio):

It can be measured as the proportion of the received packets by the destination nodes to the packets sent by the source node. This assesses the capacity of the protocol to convey data packets to the destination in the vicinity of malicious nodes.

$$PDR = (\text{received packets} / \text{packets sent}) * 100$$

Energy Consumption:

It gives the energy utilized by the node as a part of the network. It diminishes with black hole attacks of the fact that the packets transmitted between the source and destination gets dropped which prompts less transmission between the nodes.

Average end-to-end delay-(E2E Delay):

The evaluation of average time taken by a packet to transmit from source to destination. This parameter gets affected by the increase in the number of intermediate mobile nodes or malicious nodes.

$$E2E\ Delay = \frac{\sum (\text{destination reached time} - \text{sender sent time})}{\sum \text{Number of connections}}$$

Throughput: Throughput is characterized as the effective information or data packets transmitted per unit time. The parameter differs specifically with the number of packets received and is inverse proportional corresponding to the end to end delay. Accordingly, these two are the integral variables for the throughput.

$$\text{Throughput} = \frac{\sum \text{received packets}}{(\text{arrived time} - \text{send time})} * \text{packet size} * \text{time} / 1000 \text{ in kbps.}$$

Several other parameters like packet transfer rate, Detection of malicious nodes and attack detection strategies are observed and compared in simulation process shown in the graphs below.

7. SIMULATION RESULTS

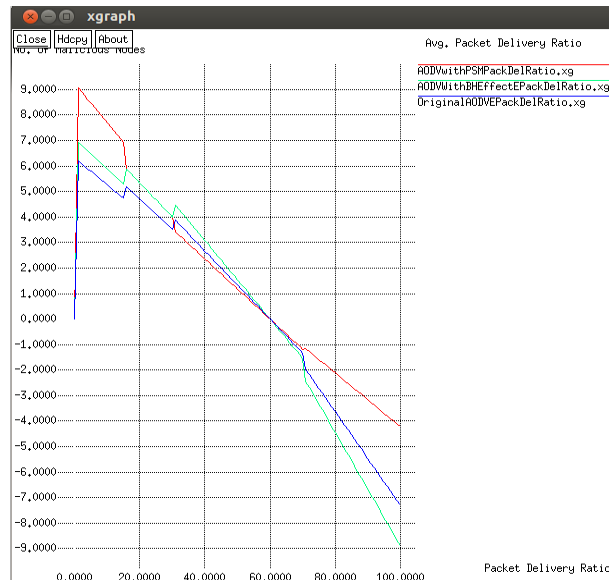


Figure 2 Average Packet Delivery ratio

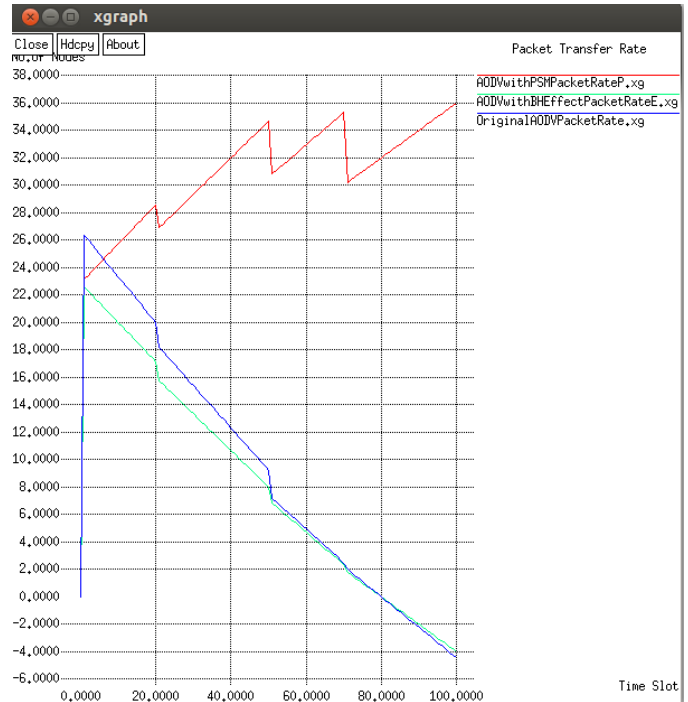


Figure 3 Packet transfer rate

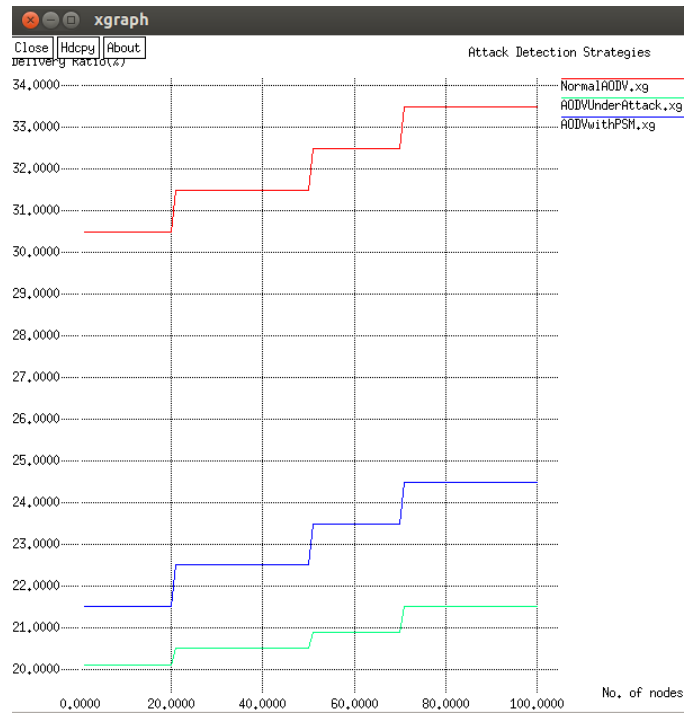


Figure 4 Attack detection strategies

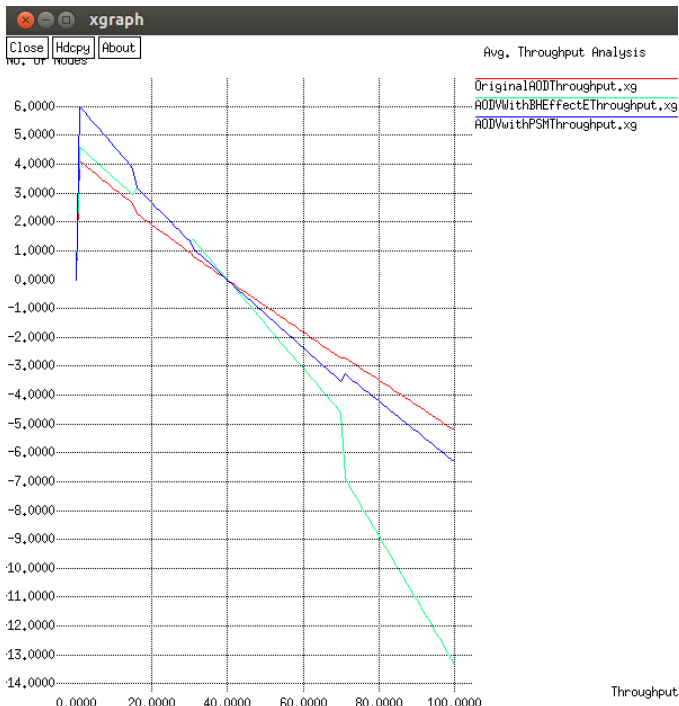


Figure 5 Average throughput analyses

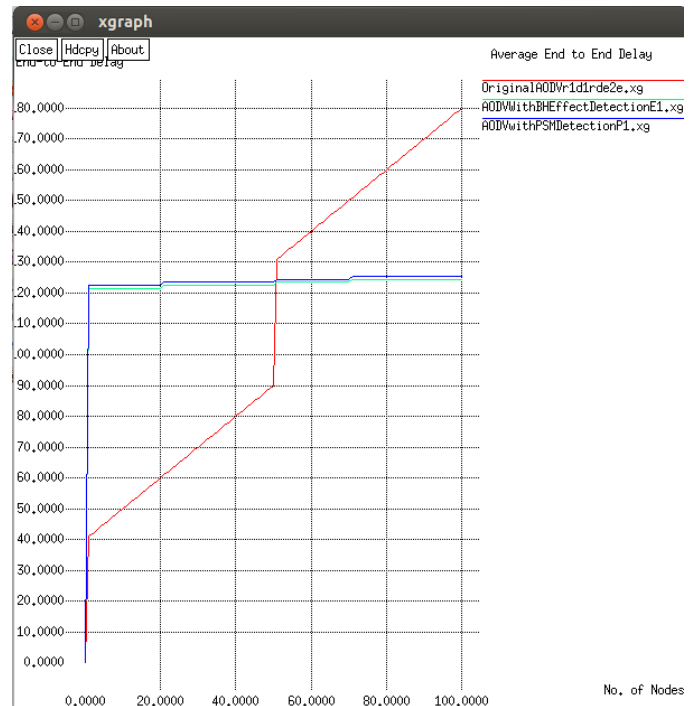


Figure 7 Average end to end delay

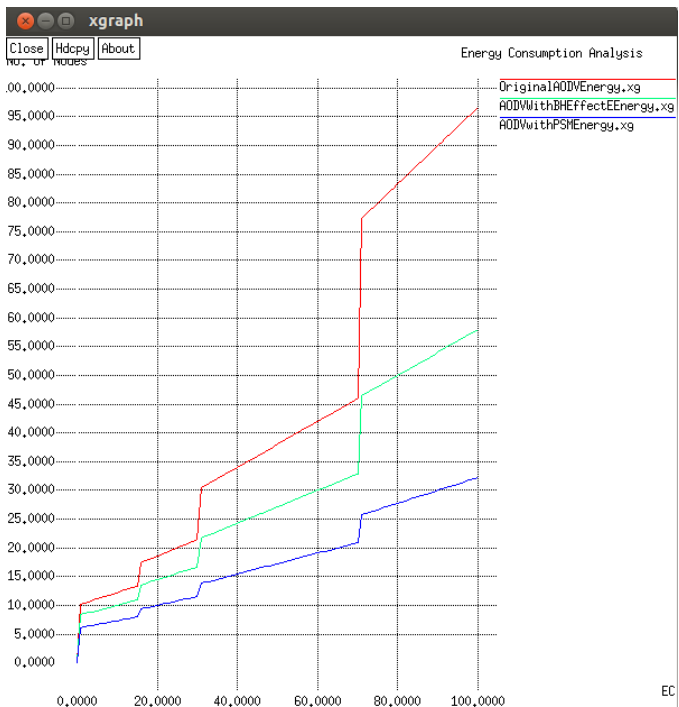


Figure 6 Energy consumption analyses

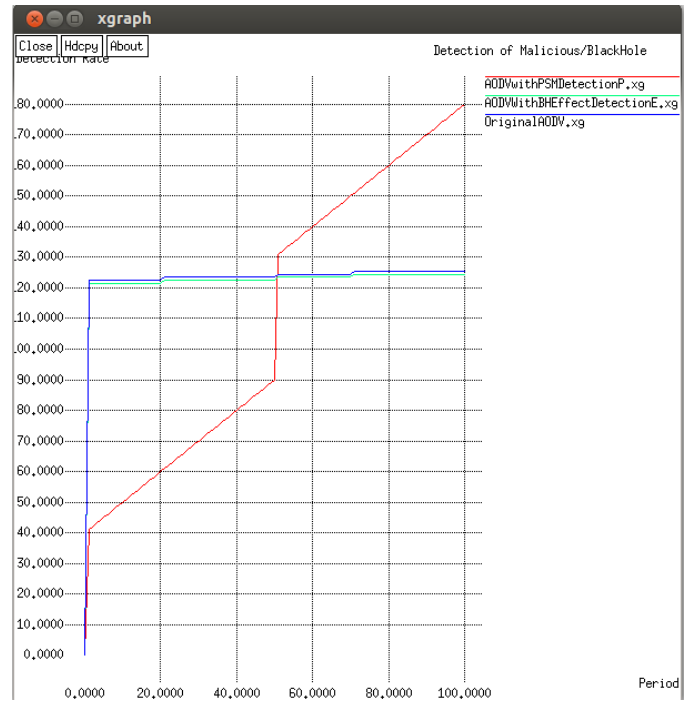


Figure 8 Detection of malicious /Blackhole nodes

8. CONCLUSION AND FUTURE WORK.

In the paper, the Proximity set method has been to demonstrate the effectiveness of performance of various parameters discussed above by implementing using NS2 and results are obtained. Compared and observed that original AODV protocol without blackhole attacks gives better

performance than AODV by using PSM to detect and remove blackhole nodes which is comparatively more than AODV with blackhole attacks. The PSM discovers a secure path for routing by preventing from malicious nodes. From the above shown simulation graphs the blackhole attacks degrades the performance of the network. In future, work can be carried out to still optimize the end to end delay and various parameters when compared with this proximity set method that may arise to increase the performance process on occurrence of black hole attack.

9. REFERENCES

1. Sumathi and B. Vinayaga Sundaram ,” An ANN Approach in Ensuring CIA Triangle using an Energy based Secured Protocol E-AODV for Enhancing the Performance in MANETS “, Indian Journal of Science and Technology, Vol 8(34), DOI: 10.17485/ijst/2015/v8i34/IPL0821, December 2015
2. Vijaya Kumar K, Dr. Somasundaram K. A Symmetric Multiple Random Keys (SMRK) Model Cryptographic Algorithm. International Journal of Innovative Research in Computer and Communication Engineering. 2015 November; 3(11), 10896-10903
3. Christeena Joseph, P. C. Kishoreraja, Radhika Baskar and M. Reji,”Performance Evaluation of MANETS under Black Hole Attack for Different Network Scenarios”, Indian Journal of Science and Technology, Vol 8(29), DOI: 10.17485/ijst/2015/v8i29/84653, November 2015.
4. V. Senthil Murugan and Dr. K. Selvakumar,” Security Measures for Black Hole Attack and Develop a Intrusion Detection System based Genetic Algorithm in Mobile Ad-Hoc Network”,Australian Journal of Basic and Applied Sciences 9(33) October 2015, Pages: 26-30
5. M. Reji, P. C. Kishore Raja, Christeena Joseph and Radhika Baskar, “ Performance Metrics of Wormhole Detection using Path Tracing Algorithm”, Indian Journal of Science and Technology, Vol 8(17), 63541, August 2015
6. Hansraj Bhakte, Prof. Rahul Kulkarni, “Prevention Of Black Hole Attacks In AODV-Based Manets Using Secure Route Discovery”, Journal of Multidisciplinary Engineering Science and Technology (JMEST) ISSN: 3159-0040 Vol. 2 Issue 7, pp 1851-1856,July 2015
7. Mr.Rahul Vasant Chavan , Prof.M S.Chaudhari, “ Enhanced DSR protocol for Detection and Removal of Selective Black Hole Attack in MANET”, International Research Journal of Engineering and Technology Volume: 02 Issue: 04 pp 510-515 July-2015.
8. Gomathi K., Dr. Parvathavarthini B., “AN EXTENSIVE ANALYSIS OFMANET ATTACKS USING SPECIAL CHARACTERISTICS”, Indian Journal of Computer Science and Engineering (IJCSE), Vol. 6 No.3 Jun-Jul 2015 pp 110-113.
9. C.V. Anchugam and K. Thangadurai,” Detection of Black Hole Attack in MobileAd-hoc Networks using Ant Colony Optimization-simulation Analysis”, Indian Journal of Science and Technology, Vol 8(13), DOI: 10.17485/ijst/2015/v8i13/58200, July 2015
10. Ranjan Bishnoi, Hardwari Lal Mandoria, “Performance Study of Black Hole Attack Detection Technique using AODV in MANET”, International Journal of Computer Science Engineering (IJCSE) Vol. 4 No.04 Jul 2015 pp 134-141.
11. V. Anand and N. Sairam ,” Methodologies for Addressing the Performance Issues of Routing in Mobile Ad hoc Networks: A Review”, Indian Journal of Science and Technology, Vol 8(15), DOI: 10.17485/ijst/2015/ v8i15/70511, July 2015.
12. Nakka Nandini, Reena Aggarwal , “ Prevention of black hole attack by different methods in MANET”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 2, February 2015 pp 297-300.
13. Y. Haripriya, K. V. Bindu Pavani, S. Lavanya and V. Madhu Viswanatham, “ A Framework for detecting Malicious Nodes in Mobile Adhoc Network”, Indian Journal of Science and Technology, Vol 8(S2), 151–155, January 2015
14. Darshana Sorathiya, Haresh Rathod,” A Review on Detection and Prevention Techniques of Wormhole Attack in MANETs”, International Journal of Science and Research (IJSR), Volume 4 Issue 1, January 2015 pp 441-444.
15. Harmanpreet Kaur, P. S. Mann,” Prevention of Black Hole Attack in MANETs Using Clustering Based DSR Protocol”, IJCST Vol. 5, Issue 4, Oct - Dec 2014 pp 278-281.
16. Santoshi K, Vijaya kumar K. An Empirical Model of Malicious Node detection and Prevention with Data rating. International Journal of Engineering Trends and Technology (IJETT). November 2014; 17(2),56-59

17. Darshana Sorathiya, Haresh Rathod, "A Review on Detection and Prevention Techniques of Wormhole Attack in MANETs", International Journal of Science and Research (IJSR), Volume 4, Issue 1, pp 441-444, January 2015.
18. Rashmi, Ameeta Seehra,"A Novel Approach for Preventing Black-Hole Attack in MANETs", International Journal of Ambient Systems and Applications (IJASA) Vol.2, No.3, September 2014 , pp 01-09
19. Akanksha Gupta, Anuj K.Gupta, "Detection and Prevention of Wormhole Attack Using Decentralized Mechanism", International Journal of Latest Trends in Engineering and Technology (IJLTET) Vol. 4 Issue 2 July 2014 pp 11-18.
20. Rahul Agarwal, Kriti Arora, Rajiv Ranjan Singh , " Comparing Various Black Hole Attack Prevention Mechanisms in MANETs", International Journal of Emerging Technology and Advanced Engineering, Volume 4, Issue 4, April 2014, pp 157-166.
21. Chetana Khetmal, Shailendra Kelkar, Nilesh Bhosale, "MANET: Black Hole Node Detection in AODV", International Journal of Computational Engineering Research Vol, 03 Issue, 6 pp79-85
22. Jaspal Kumar, M. Kulkarni, Daya Gupta , " Effect of Black Hole Attack on MANET Routing Protocols", J. Computer Network and Information Security, 2013, 5, 64-72 Published Online April 2013 in MECS DOI: 10.5815/ijcnis.2013.05.08.
23. Vipin Khandelwal, Dinesh Goyal, " BlackHole Attack and Detection Method for AODV Routing Protocol in MANETs" , International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013, pp 1555-1559
24. Romina Sharma, Rajesh Shrivastava, "Modified AODV Protocol To Prevent Black Hole Attack in Mobile Ad-hoc Network", INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT, April, 2013 Vol 2 Issue 4, pp 476-487.
25. Tanu Preet Singh, Prof. R.K Singh, Jayant Vats and Manmeet Kaur," Optimized Black Hole Detection and Prevention Algorithm (OBHDPA) for Mobile Ad Hoc Networks", International Conference on Computer Science and Information Technology (ICCSIT'2011) Dec. 2011 pp 57-60.
26. Satoshi Kurosawa, Hidehisa Nakayama, Nei Kato, Abbas Jamalipour and Yoshiaki Nemoto," Detecting Blackhole Attack on AODV-based Mobile Ad Hoc Networks by Dynamic Learning Method", International Journal of Network Security, Vol.5, No.3, PP.338-346, Nov. 2007
27. Teerawat Issariyakul , Ekram Hossain," Introduction to Network Simulator NS2",Springer, DOI: 10.1007/978-0-387-71760-9.
28. C SivaRam Murthy, B S Manoj , "Adhoc Wireless Networks" , Protocols and Architecture, Pearson publishers.

A Greedy Approach To Out-Door WLAN Coverage Planning

Gilbert M. Gilbert,
College of Informatics and Virtual Education,
The University of Dodoma

Abstract— Planning for optimal out-door wireless network coverage is one of the core issues in network design. This paper considers coverage problem in outdoor-wireless networks design with the main objective of proposing methods that offer near-optimal coverage. The study makes use of the greedy algorithms and some specified criteria (field strength) to find minimum number of base stations and access points that can be activated to provide maximum services (coverage) to a specified number of users. Various wireless network coverage planning scenarios were considered to an imaginary town subdivided into areas and a comprehensive comparison among them was done to offer desired network coverage that meet the objective.

Keywords— **greedy algorithms, outdoor-wlan, coverage planning, greedy algorithms, path loss.**

I. INTRODUCTION

The use of wireless networks has drastically evolved to change the way the world communicates. Nowadays users and devices are turning toward the use of wireless communication technologies because of advantages such as easy and simple communication. Moreover, wireless communication technologies allow users and devices to conveniently access their applications and freely move from one point to another in a given area.

However, owing to their seamlessly advantages, wireless networks are placed with such a huge demand by users and are therefore becoming more complex both in the design and maintenance. Also there are constraints such as a limited number of places to position masts, legal and physical on the limits of powers and frequencies on the location, traffic demand of users and economic situation [1].

Providing continuous, reliable, fast and low cost services are what competitors are trying to achieve with the design of wireless networks. As a result, robust and reliable wireless network designs are needed to meet users' ever-rising demands. Proper planning and implementations of wireless networks are therefore very important to ensure that they are cost-effective to service providers and also do not compromise the quality of service to users. To the service providers, the placement of wireless access points or base stations is crucial. These telecommunication devices need to be adjusted, arranged and sometimes re-arranged to provide optimal coverage in a given geographical area. From the economic point of view, the process of optimisation should be of low

cost with best outcomes in terms of profit.

Therefore, this paper considers the planning and design of out-door wireless networks with the goal of achieving as maximum coverage as possible using greedy algorithm. This paper is split into five sections. Sections II, III and IV include literature review, methodology and results and discussion respectively. The last section includes concluding remarks.

II. LITERATURE REVIEW AND RELATED WORKS

The following paragraphs provides an overview of various approaches in the design and optimisation of wireless networks.

A. Cell and Coverage Planning

Cell and coverage planning problems are the cores issues in the design of wireless network. They have to address fundamental requirements such as network traffic, the topography of the area, the specification of propagation and system desired capacity. This area has attracted a number of research works with examples being [2-4].

B. Manual Design and Optimization

In most cases wireless network designs are done manually, either on paper, visiting the building or plots. Experienced engineers use their knowledge to constrain the number of parameters during the planning process. However, using these rule-of-thumbs in the designs can produce a single plan that, of course, is potentially feasible solution but lacks alternative plans to evaluate the cost and benefit of each plan as well as pose failure problems such as site acquisition problems which introduce heavy task of iteration of the plan which may cause degenerative plan and make the process very costly [5].

C. Automatic Design and Optimisation Techniques

1) AP/ Base station Placement and Selection

The problem of minimizing the number of candidate APs that are able to cover all users is a well-known combinatorial optimisation problem as studied by [6] with proposed optimization techniques. [7] came up with classical approach for coverage planning based on random search heuristics. The placement of base stations then is crucial to the effectiveness and efficiency of the network in design. The positioning of the base stations has to match the network requirements and objectives as having too few base stations risk the loss of service to users (unable to obtain any service) with too many

users trying to access the same single base station. Moreover, having too many base stations may create excess interference [7] between cells with a similar frequency (co-channel interference) or adjacent channel interference as well as incurring huge cost of installing and maintaining the equipment.

2) Idealized Physical Layout of Wireless Network

In the idealized model, each cell has the shape of a circle and perfect spherical radio propagation is assumed. The signal coverage of the base stations may overlap allowing the user to listen the signals from more than one BS. For example, a user in region H can listen to the signals from BSs 0, 5 and 6. The user then is in a good position to choose the strongest signal from among the BS.

To be able to access the network, the user's device (or terminal) needs to transmit or receive the radio signal of a base station at a reasonable amount of field strength which is in an adequate level.

One of the simple ways of planning wireless network coverage is to have a set of possible positions of user terminals (Test Point, TPs), in the service area and a set of base station candidate sites (CS). Then a subset of candidate sites will have to be picked up so as to provide some network requirements such as coverage and capacity to all TPs.

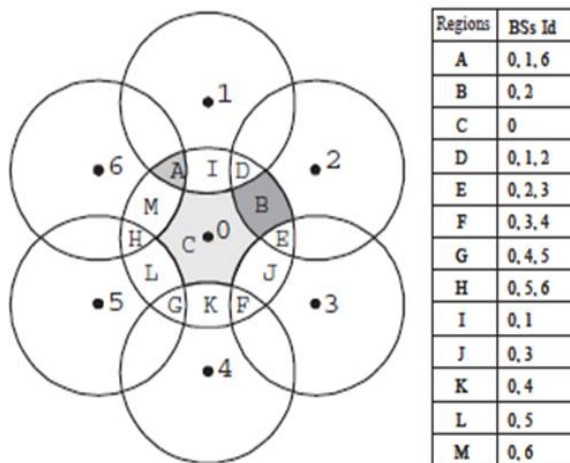


Fig 1 A layout of wireless network in idealized model

3) Power Control

Various studies have also looked into adjusting the transmission range, power control, for design a topology satisfying a given constraint [9]. Other researchers tried to create a topology that is driven by QoS using power control. Thompson et al [10] proposed techniques of optimization that focus on energy consumption of wireless networks by identifying major parts where power is consumed a lot and also suggested ways to reduce overall long-term energy consumption by decreasing transceiver power and upgrading hardware.

D. Optimization Algorithms

There are various approaches to coverage optimisation algorithms for wireless systems in the literature. In many

works the focus has been on gradient descent [11] random search [12] and genetic algorithms [13]. Gradient algorithm seems to attract the attention because there are ways to prove convergence and therefore guarantees convergence under specific set of conditions [14].

Genetic algorithms as explained by [15] have been widely used in wireless network design particularly in cell planning [16].

Hill climbing algorithms tend to generate a sequence of solutions to the network design problem [17]. The downside of these algorithms is that they tend to offer local rather than global optima.

Simulated annealing as researched by [18] are also used as one of the cell planning approaches with notable examples from [8,19].

Another important algorithm in the optimization realm is Tabu search (TS). TS has been used to solve cell planning problem as in [7,20].

E. Greedy Algorithms

These algorithms are best in providing the initial design for further development. The algorithms follow the problem-solving approach of heuristic algorithms by making the locally optimum choice at each stage of evaluation in the hope of finding the global optimum solution [21]. The algorithm works by evaluating each local optimal solution at each looping stage, and usually without considering future consequences. This algorithm will not find a global optimal solution, but finds sub-optimal answers quickly.

These algorithms have been used in various notable researches in cellular networks, such as [22] in cell planning, and [23] for channel assignment in cellular networks.

III. METHODOLOGY

A. Data for Experiment and Simulation

A collection of data that was widely used by a number of researchers such as [24] was also employed in this work.

B. Generic Network for Problem Description

A model proposed by (Reininger and Caminada 1998) was adopted in which a network is defined within a working area P. Any point in P is defined by its Cartesian coordinates (x, y). Points within P where propagation and service information is available are known as Service Test Points (STP) represented by

$$S = \{S_1, S_2, \dots, S_{n_s}\}$$

where n_s is the total number of service test points in P.

A typical STP is represented by S_i with coordinate (x_i, y_i) where $1 \leq i \leq n_s$ and mesh data provided for each STP include

- Propagation loss estimates
- Service threshold requirements

Other, called engineering data, needed is a list of candidate sites, where BS could potentially be located.

1) Service Requirements

A network provides a service based upon criteria specified by the network operator. The nature of the expected service for a network defined on P is given by S, and the values associated with the points in S give the service threshold (in dBm) for the required service. The service requirement at an STP S_i can be represented by S_{qi} .

2) Propagation Losses

The propagation loss from every candidate site to each STP defined on P is given by Q, where

$$Q = \{Q_1, Q_2, \dots, Q_{n_{sites}}\}.$$

Q_i contains estimates of propagation losses to each STP defined on P, that is, Q_i contains n_s propagation loss estimates (in dB) and $Q_{n_{sites}}$ is total number of candidate sites.

3) A Base Station

The model also considers a set of base stations, B, such that:

$$B = \{B_1, B_2, \dots, B_{n_B}\}$$

where n_B is the total number of base stations.

The base station would have a number of operational parameters to be configured. In this research only transmitting power (in dBm) represented as BT_{sj} is considered.

4) Received Power

For a given base station B_j , the received field strength (power) at a service point S_i is given by

$$P_j(S_i) = BT_{sj} - Q_j(S_i) \quad (1)$$

The power of a base station is given in dBm

5) Cell

A cell C_j is defined for base station B_j as the set of points such that:

$$C_j = \{S_i: P_j(S_i) \geq S_{qi} \text{ and } P_j(S_i) > P_k(S_i) \forall k, 1 \leq k \leq n_B, k \neq j\}$$

i.e., for all points in cell C_j , B_j is the best server (provides the strongest signal)

C. Design Objective

The design objective of the wireless network planning is to maximize coverage while minimizing number of base stations. The coverage objective can be modelled mathematically. It requires all service test points receive at least one signal above its threshold, i.e., for all service test points S_i , it is required

$$\sum_{i=1}^{n_B} \mu_{ij} = n_s \quad (2)$$

$$\text{where } \mu_{ij} = \begin{cases} 1 & \text{if } \dots S_i \in C_j \\ 0 & \text{Otherwise} \end{cases}.$$

The point S_i is said to be covered if it receives at least one signal above its service threshold in C_j by base station B_j .

1) Objective Function and Constraints

The objective function of the problem tries to minimize the number of base stations required as well as to maximize received signal strength (coverage) over all reception sites. Network coverage (Z_1) as a network cost is taken as a ratio of number of service test points covered to the total number of service points.

$$Z_1 = \frac{n_{covered}}{n_s} \quad (3)$$

Now the objective function is to minimize number of base stations (number of subsets in B) such that their union covers all points in S.

$$Z_2 = \min \sum_{j \in I} B_j \quad (4)$$

Subject to

$$\sum_{i=1}^{n_B} \mu_{ij} = n_s$$

$$P_j(S_i) - S_{qi} > 0, 0 < Z_1 \leq 1,$$

i is the set of possible base station positions, $i = \{0, \dots, n_B\}$.

D. Algorithm Design

1. FOR each base station B_j in the base station set B
 - a. Base station B_j status = 0;
2. End FOR
3. Turn on base station 1
4. Calculate network coverage, *Old_Coverage* equation (3)
5. FOR base station $j=2$ to number of base stations, n_B
 - a. Base station status = 1;
 - b. Calculate new network coverage, *New_Coverage* (equation 3)
 - c. IF *New_Coverage* > *Old_Coverage*
 - i. Base station status = 1
 - ii. Set *Old_Coverage* = *New_Coverage*
 - d. ELSE
 - i. Set Base station status = 0
 - e. END IF
6. END FOR

The algorithm evaluates the wireless network design by using its network coverage values. It starts by setting off all the base stations (in line 1(a)). The first base station is added (activated or switched on) to the network, in line 3. The cost of the network, that is, network coverage, *Old_Coverage* is then calculated using the added base station parameters. Then other base stations are added one by one, line 5. A new value of network coverage, *New_Coverage* will be calculated each time a base station is added. If the new value of the network coverage is much better than the previous value (improves the coverage of the network) after the addition of that base station, then the base station is kept on, otherwise it is removed (deactivated) or switched off, and the algorithm moves to the

next base station and repeating the same procedure until the last base station.

E. The Greedy Algorithm Implementation

The greedy algorithm decides which base stations are best to switch on. The algorithm loops through the base stations, which are all initially off, switching them one at a time. Each time a base station is switched on, the algorithm calculates the new solution value (network coverage ratio) with the base stations that are on as shown in Figure 2.

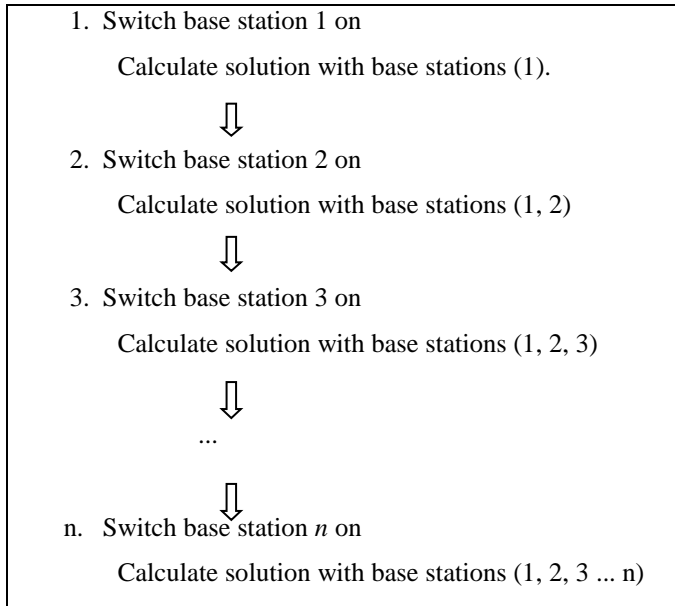


Fig 2:

			3,14,15,1 6,17,18	
-60	40	0.951	1,2,3,4,5, 6,7,8,9,1 0,11,12,1 3,14,15,1 6,17,18	5121
-60	50	0.997	1,2,3,4,5, 6,7,8,9,1 0,11,12,1 3,15,16,1 7,18	5366
-60	60	0.999	1,2,3,4,5, 6,7,8,9,1 0,12,13,1 5,16,17,1 8	5378

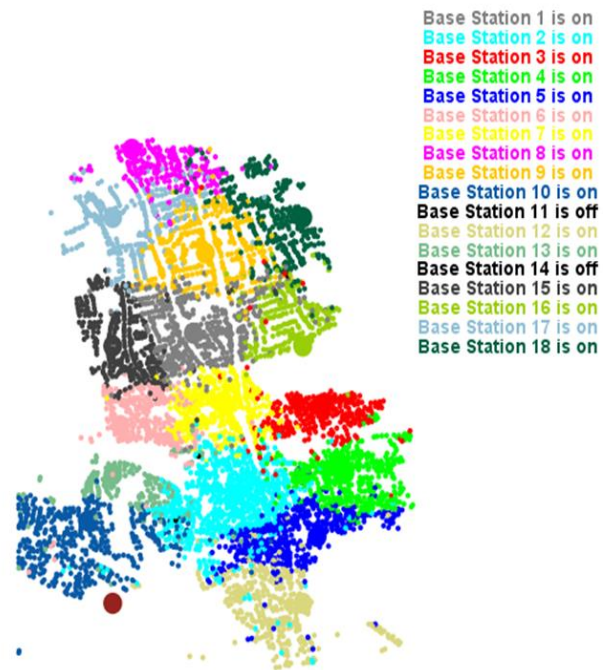


Fig 3: Coverage planning when threshold is at -60dBm and BS power is at 60dBm

IV. RESULTS AND DISCUSSION

The following paragraphs depict the results of different scenarios.

A. Scenario 1: When Threshold is Constant

1) Threshold at -60dBm

From Table 1, the maximum coverage occurs when threshold was -60dBm and BS signal strength at 60dBm, and two base stations (11 and 14) were off. Figure 3 depicts the coverage at 60dBm thresholds and 60dBm base station power.

Table 1: Results of Simulation when Threshold is constant at -60dBm

Thresho Id (dBm)	BS Signal Strength (dBm)	Network Coverag e ratio	Base Station ON	Users Covered
-60	20	0.052	1,2,3,4,5, 6,7,8,9,1 0,12,13,1 4,15,16,1 7,18	282
-60	30	0.393	1,2,3,4,5, 6,7,8,9,1 0,11,12,1	2114

2) Threshold at -70dBm

Table 2: Results of simulation when threshold is -70dBm

Threshol d	Base Station Power	Network Coverage Ratio	Base Station ON	Users Covere d
-70	20	0.393	1,2,3,4,5, 6,7,8,9,10 ,11,12,13, 14,15,16, 17,18	2114
-70	30	0.952	1,2,3,4,5, 6,7,8,9,10 ,11,12,13,	5121

			14,15,16, 17,18	
-70	40	0.997	1,2,3,4,5, 6,7,8,9,10, ,11,12,13, 15,16,17, 18	5366
-70	50	0.999	1,2,3,4,5, 6,7,8,9,10, ,12,13,15, 16,17,18	5378
-70	60	1	1,2,3,4,5, 6,7,8,9,10, ,11,15,17	5381

From Table 2, there is a significant change of network coverage ratio between the first row (BS power at 20dBm) and the second row (BS power at 30dBm) from the 0.393 to 0.952 with all the base stations switched on. However, the network ration of 1 is obtained when base station power was set to 60dBm with five base stations switched off. Figure 4 illustrates the scenario when BS power is at 20 in which all the base stations are switched, and Figure 5 shows the coverage planning for this scenario.

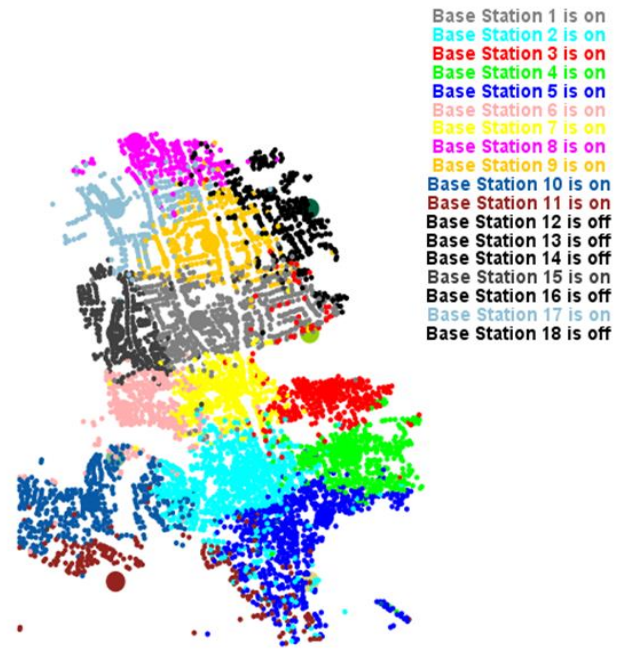


Fig 5: Coverage planning when threshold is at -70 dBm and BS power is at 20dBm.

3) *Threshold at -100 dBm*

Table 3: Results of simulations when threshold is at -100dBm

Threshold	Base Station Power	Network Coverage Ratio	Base Station ON	Users Covered
-100	20	0.999	1,2,3,4,5, 6,7,8,9,10, ,12,13,15, 16,17,18	5378
-100	30	1	1,2,3,4,5, 6,7,8,9,10, ,11,15,17	5381
-100	40	1	1,2,3,4,8, 9	5381
-100	50	1	1,2	5381
-100	60	1	1,2	5381
-100	80	1	1	5381

It can be seen from Table 3 that, the maximum network coverage ratio of 1 is obtained when base stations power is from 30 to 80 dBm with the best case scenario at -100 dBm threshold and 80dBm as base station power.

B. *Scenario 2: When base station power is constant*

1) *Base stations power at 40 bBm*

Table 4: Results of simulations when BS power 40dBm

Threshold	Base Station Power	Network Coverage Ratio	Base Station ON	Users Covered
-60	40	0.951	1,2,3,4,5, 6,7,8,9,10, ,11,12,13,	5121

Base Station 1 is on
Base Station 2 is on
Base Station 3 is on
Base Station 4 is on
Base Station 5 is on
Base Station 6 is on
Base Station 7 is on
Base Station 8 is on
Base Station 9 is on
Base Station 10 is on
Base Station 11 is on
Base Station 12 is on
Base Station 13 is on
Base Station 14 is on
Base Station 15 is on
Base Station 16 is on
Base Station 17 is on
Base Station 18 is on



Fig 4: Coverage planning when threshold is at -70 dBm and BS power is at 20dBm

			14,15,16, 17,18	
-70	40	0.997	1,2,3,4,5, 6,7,8,9,10, 11,12,13, 15,16,17, 18	5366
-80	40	0.999	1,2,3,4,5, 6,7,8,9,10, 12,13,15, 16,17,18	5378
-90	40	1	1,2,3,4,5, 6,7,8,9,10, 11,15,17	5381
-100	40	1	1,2,3,4,8, 9	5381

In the table 3, the maximum network coverage ratios are 1 and were obtained when thresholds were at -90 and -100 dBm. At threshold of -90 dBm, five base stations are off, while at -100 dBm, only six base stations are switched on. Figure 6 demonstrate the network coverage at -100 dBm as threshold.

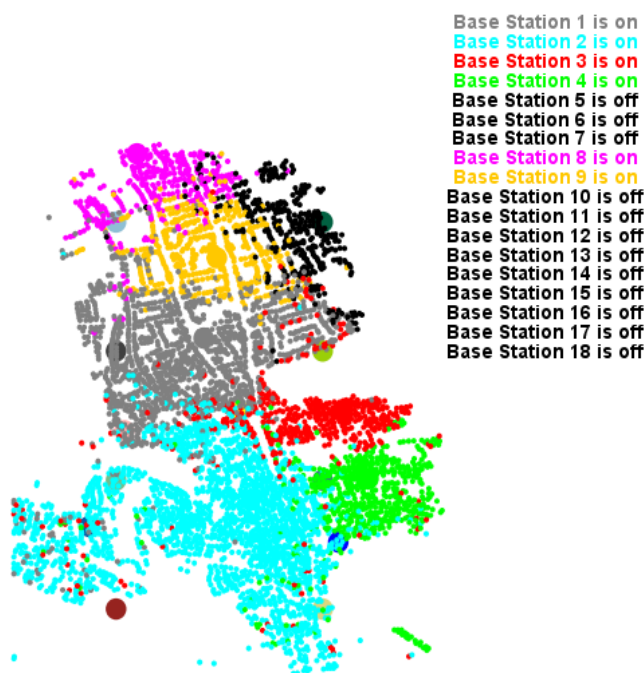


Fig 6: Coverage planning when threshold is at -100 dBm and BS power is at 20dBm

On interpreting the wireless network design simulations, it can be seen clearly how the users of the wireless network are distributed over the area. It can be easily observed how the numbers of users that connect to a particular base station vary as thresholds and base station power are changed. The black dots in the figures represent users who are not connected to any base station, which can be attributed by the fact that they don't receive enough signals either because of the high threshold values or less power emitted from the base stations. In the scenarios in which the network coverage ratio is 1, they

suggest that the network design objective has been met, that is good coverage.

If figure 4 is examined, it can be seen that all base stations have been switched on, but the coverage is not fully. This can be translated into prioritization of service provision and delivery depending on the policy of the service providers. Moreover, that could also mean that the service provider engineers should continue adjusting network parameters to provide full network coverage.

V. CONCLUSION

The algorithm implemented in this work is simple and it only takes into account one constraint. Nevertheless, the algorithm performs quickly and provides good sub-optimal solutions in as seen from results. When it comes to designing a fully-operational network other factors such as capacity of base stations, traffic demand and interference would also have to be considered. In that kind of environment, the algorithm accuracy and performance can be fully tested. In addition, from the results it can be seen that if the user wishes to design a wireless network that involves large number of base stations (catering for redundancy, for example) then low uniform powers should be assigned to base stations.

REFERENCES

- [1] Ryan, D. et al. 2005. 3.5 GHz Broadband Fixed Wireless Network Design for Rural Deployment. 14th IST Mobile & Wireless Communications Summit. Dresden.
- [2] S.M. Allen, S. Hurley and R.M. Whitaker. 2001., Spectrally Efficient Cell Planning in Mobile Wireless Networks", IEEE, VTC'01, pp-931-935.
- [3] S.M. Allen, S. Hurley, R.K. Taplin, A.A. Wade. 2003, Cost Effective Cell Planning of Millimeter Broadband Systems", IEEE, pp-280-284, Michael Faraday House, Six Hill way. Stevenage SG1.
- [4] Young Ha Hwang, Sung-Kee Noh, and Sang-Ha Kim (2006), "Determination of Optimal Cell Capacity for Initial Cell Planning in Wireless Cellular Networks", KIPS (ISSN: 1738-8899).
- [5] Hurley, S. et al. 2004. Smart Cell Planning and Optimization for UMTS. 5th IEEE International Conference on 3G Mobile Communication Technologies, pp 34-38.
- [6] Lee, C. and Kang, H. 2000. Cell planning with capacity expansion in mobile communications: A tabu search approach. IEEE Transactions on Vehicular Technology. 49(5), pp.1678-1691, March 2000.
- [7] Hurley, S. (2002) Planning Effective Cellular Mobile Radio Networks, IEEE Transactions on Vehicular Technology 51(2), pp. 243-253, ISSN 0018-9545.
- [8] Hu, L. 1993. Topology control for multihop packet radio networks. IEEE Transactions on Communications.
- [9] Thompson, J. et al. 2010. Base Station Location for Minimal Energy Consumption in Wireless Networks. IEEE 73rd Vehicular Technology Conference, pp. 1-5.
- [10] Rappaport, T. 2001. Wireless Communications Principles and Practice. 2nd ed. Prentice Hall
- [11] Anderson H. and McGeean, J. 1994. Optimizing microcell base station locations using simulated annealing techniques. Proceedings of WC. Stockholm, Sweden, June 1994, vol. 2, pp. 858-862.
- [12] Lieska, K., E. Laitinen, E and Liihteenmiki, J. 1998. Radio coverage optimization with genetic algorithms. Proceedings of PIMRC, Boston, MA, September 1998, vol. 1, pp. 318-322.

- [13] Kamenetsky, M. and Unbehaun, M. 2002. Coverage Planning for Outdoor Wireless LAN Systems. International Zurich Seminar on Broadband Communications
- [14] Schmitt, L. 2001. Theory of Genetic Algorithms. Theoretical Computer Science 259 (1-2), pp. 1-61.
- [15] Valenzuela, S. Hurley, and D.H. Smith. A permutation based genetic algorithm for minimum span frequency assignment. In Parallel Problem Solving from Nature –PPSN V: 5th International Conference, volume 1498 of Lecture Notes in Computer Science, pages 907–916. Springer Verlag, 1998.
- [16] Aarts, E. and Lenstra, J. 1997. Local search in combinatorial optimization. John Wiley, Chichester.
- [17] Aarts, E. and Korst, J. 1989. Simulated Annealing and Boltzmann Machines. John Wiley and Sons.
- [18] Anderson, H. and McGeehan, J. 1994. Optimizing microcell base station locations using simulated annealing techniques. In Proceedings 44th IEEE Conference on Vehicular Technology, pp 858–862
- [19] Han, J. et al. 2001. Genetic approach with a new representation base station placement in mobile communications. In Proceedings 54th IEEE Conference on Vehicular Technology, vol.4, pp. 2703–2707.
- [20] Sahni, S. 2000. Data Structures, Algorithms, and Application in Java. Singapore: McGraw-Hill Book Co
- [21] Tutschku, K. 1998. Interference minimization using automatic design of cellular communication networks. In Proceedings of the IEEE VTC'98 Conference, pp. 634–638.
- [22] Battiti, R., Bertossi, A. and Cavallaro, D. 2001. A randomized saturation degree heuristic for channel assignment in cellular radio networks. IEEE Transactions on Vehicular Technology, 50(2), pp. 364–374.
- [23] P. Reininger and A. Caminada, “Model for GSM radio network optimization,” presented at the 2nd ACM Int. Conf. Discrete Algorithms and Methods for Mobility, Dallas, TX, 1998.

Cerebellar model articulation controller network for segmentation of computer tomography lung image

¹Benita K.J Veronica, ²Purushothaman S., Rajeswari P.,

¹ Research Scholar, Mother Teresa Women's University, Kodaikanal, India.	² Associate Professor, Institute of Technology, Haramaya University, Ethiopia.	
-------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------	--

Abstract-This paper presents the implementation of CMAC network for segmentation of computed tomography lung slice. Representative features are extracted from the slice to train the CMAC algorithm. At the end of training, the final weights are stored in the database. During the testing the CMAC, a lung slice is presented to obtain the segmented image.

Keywords: CMAC; segmentation; computed tomography; lung slice

1. Introduction

Lung nodules in human bodies indicate the lung abnormalities. A pulmonary nodule is the most common manifestation of lung cancer. Lung nodules are approximately spherical regions of relatively high density that are visible in X-ray images of the lung. Large malignant nodules with a diameter greater than 1 centimeter can be detected using the traditional imaging equipment easily. Such large nodules can be diagnosed by needle biopsy or bronchoscopy techniques. Imaging techniques are required for identifying small nodules.

Nodules can have any shape with solid, transparent, soft or hard. Due to all these factors, automatic lung nodule detection from the CT images is still a challenging issue. Moreover, some mathematical algorithms have been proposed in the past. They must be supported by a good radiologist to make effective decisions.

Jaspinder et al., 2014, proposed a computer aided diagnostics (CAD) system to segment the lung tumor using region of interest (ROI) and gray level co-occurrence matrix (GLCM).

Amjed et al., 2014, proposed a CAD system with segmentation by thresholding and labeling. Bhavanishankar and Sudhamani, 2015, provided a survey of different techniques to classify the lung nodules. Ramya et al., 2015, worked on classification of lung nodules using adaptive graph patch based division and developed feature set.

Suresh Tripathi and Xuqiu Zhen, 2015, discussed the characteristics of solitary pulmonary nodule (SPN).

Sudha.V, Jayashree.P, 2012, developed an efficient lung nodule detection system by performing nodule

segmentation through thresholding and morphological operations. Regmi Meghna, and Wang Pei Jun, 2015, reviewed about the CT appearances of solitary pulmonary nodules.

II Methods

A. Contextual clustering (CC) for feature extraction

The training (Phase-1) and testing (Phase-2) of CMAC network requires features from the lung slice. In both the Phase-1 and Phase-2, features such as the mean of the (3X 3 window pixels) 9 intensity values, the summation of 9 intensity values and the TH_{cc} obtained from CC algorithm for a moving overlapping window are given as inputs to the input layer of the CMAC topology. The CC segments a data into category 1 (a_0) and category 2 (a_1). They are assumed to be drawn from the standard normal distribution. The following steps are adopted for implementing contextual clustering.

Step 1: The decision parameter T (positive) and weight of neighborhood information β (positive) are defined. The neighborhood can be denoted by N_n . The data is denoted by D_i .

Step 2: The data is classified when $D_i > T\alpha$ to a_1 and data to a_0 . The classification is stored to G_0 and G_1 .

Step 3: For each data 'i', the number of data u_i is counted, belonging to class ω_1 in the neighborhood of data 'i'. The data outside the range belong to a_0 .

Step 4: The data with is assigned to a_1 and other data to a_0 . The classification is stored to variable C_2 .

Step 5: If $G_2 \neq G_1$ and $G_2 \neq G_0$, copying of G_1 to G_0 , and G_2 to G_1 are done and returned to step 3, otherwise the process is stopped and returned to G_2 .

The contextual clustering is a supervised algorithm. An overlapping window of size (3 X 3) is used to scan the image for segmentation. Segmentation of image depends upon 1) a defined threshold value ($T_h=140$) by the user which is based on the histogram of the original image, 2) a moderating parameter β_i which is in the range of $0 < \beta_i < 0.5$, 3) the median value, 4) the total number of intensity values (v) > 10 inside the window, excluding the already identified median value.

Input: CT Lung image is input to CC algorithm.

Step 1: The intensity values are sorted.

Step 2: The median of the intensity values, C_m , is obtained.

Step 3: The number of intensity values greater than 1 is found.

Output: CC values are obtained using the formula.

$$TH_{cc} = \text{Median value} + \frac{\beta i}{\text{threshold}} \times \left\{ v - \frac{\text{Window size}}{2} \right\} \quad (1)$$

The contextual value TH_{cc} is calculated by the equation 1. The value is compared with a set threshold (Threshold $T_h=160$). If the contextual value is less than or equal to the set threshold, then 0 is assigned to the center of window else, 255 is assigned to the center of the window. The presence of 255 in a region in the image shows the segmented object.

B. Cerebellar model articulation controller (CMAC) neural network

An associative memory algorithm called CMAC artificial neural network is used for slice lung segmentation. Quantization is performed on a pattern which is presented in the input layer of the CMAC. In Figure 1 a CMAC architecture is presented. In the training and the testing process, a quantized pattern is presented in the input layer. A searching process is used in quantized space to obtain a different value that is used as input pattern for subsequent processing. Linear processing obtains required outputs.

Steps involved in implementing CMAC:

Step 1: Present a pattern.

Step 2: Each pattern is quantized into many ranges. A value of 0.2 is quantized into (0.05 -0.1), (>0.1 till 0.19), (>0.19 till <0.26). Now, different quantization ranges are created. The given value 0.2 can be associate to the third range. If a value is 0.11, then the representation can be [0, 1, 0]. Similarly, if the value is 0.06, then it can be associate with [0, 0, 1]. Also, a third value can be associated with equal representation. The converted input patterns is [0, 0, 1, 1, 0, 0, 0, 1, 0]. Linear processing is performed on this pattern as per the topology of the CMAC.

Step 3: In the testing process, pattern from a new image is converted into binary representation as defined in Step 2. The output in the output layer is obtained. The output is further compared with a template for the required result.

III. Results and discussions

Images from lung image database consortium (LIDC) are used to evaluate the segmentation accuracy of the proposed CC / CMAC. LIDC has been used exclusively for segmentation. The LIDC Database contains a total of 1018 helical thoracic CT scans. The 1018 CT scans had been acquired from 1010 different patients.

A. Precision –recall analysis

Precision is a measure of result relevancy. Measure of truly relevant results returned is a Recall. The curve represents both high recall and high precision. Low false positive rate represents high precision. Low false negative rate represents high recall High scores for both presents results (high precision) accurate, plus a majority of all positive results (high recall) the classifier is returning.

A system gives various results with high recall and low precision values, Compared to the training labels. Most of its predicted labels are not correct. It is just the opposite of a system returning few results with high precision and low recall, We see the predicted labels are correct compared to the training labels. An ideal system with high precision and high recall will return many results, with all results labeled correctly.

$$P = T_p / (T_p + F_p) \quad (2)$$

$$R = T_p / (T_p + F_N) \quad (3)$$

F_1 score represents harmonic mean (precision-recall)

$$F1 = 2 \{ (P \times R) / (P + R) \} \quad (4)$$

Where T_p is the true positive in which a present segmentation is perfect. F_p is the False Positive in which unwanted is incorrectly segmented. F_N is the False Negative in which actual area is incorrectly segmented and T_N is the True Negative in which a unwanted area is correctly segmented.

Figure 2 indicates that for CMAC, as the recall is more precision is also more. In case of CC, as the recall is more precision is less. The harmonic mean of CC (green color) is less than that of CMAC (black). This indicates that CMAC performs well in segmentation when compared to that of CC.

IV. Conclusions

This paper presents the implementation of CMAC network for segmentation of LIDC lung slice. Representative features are extracted from the CC algorithm to present as inputs to the CMAC algorithm. The input values are further quantized to linearly process in the forward layer of the CMAC. The CMAC performance in lung segmentation is compared with that of CC and found that CMAC performs better segmentation.

References

- [1] Amjed S. Al-Fahoum, Eslam B. Jaber, and Mohammed A. Al-Jarrah, 2014, Automated detection of lung cancer using statistical and morphological image processing techniques, Journal of Biomedical Graphics and Computing, Vol. 4, No. 2, pp.33-42.
- [2] Aswathy S. Nair, Jisu Elsa Jacob, 2015, Automatic Lung Nodule Detection on CT Image Using Region Growing,

- International Journal of Engineering and Advanced Technology, Vol.4, Issue-5, PP.157-159.
- [3] Bhavanishankar .K and M.V.Sudhamani, 2015, Techniques for detection of solitary pulmonary nodules in human lung and their classifications -a survey, International Journal on Cybernetics and Informatics, Vol.4, pp.27-40.
- [4] Jaspinder Kaur, Nidhi Garg, and Daljeet Kaur, 2014, Segmentation and Feature Extraction of Lung Region for the Early Detection of Lung Tumor, International Journal of Science and Research, Vol.3, Issue 6, pp.2327-2330.
- [5] Ramya Preethi S., Vijayalakshmi R., and Deepa P., 2015, Lung Nodule Detection Based on Semi Supervised Classification, International Journal of Innovative Research in Science, Engineering and Technology, Vol.4, Issue 6, pp.250-254.
- [6] Regmi Meghna, and Wang Pei Jun, 2015, CT appearance of solitary pulmonary nodules; differentiating benign and malignant: a review, Biomedical letters, Vol.1, Issue 2, pp.74-80.
- [7] Sudha V., and Jayashree P., 2012, Lung Nodule Detection in CT Images Using Thresholding and Morphological Operations, International Journal of Emerging Science and Engineering, Vol.1, Issue-2, pp.17-21.
- [8] Suresh Tripathi and Xuqiu Zhen, 2015, Differentiation of Benign and Malignant Solitary Pulmonary Nodule: Literature Review, Advances in Lung Cancer, Vol.4, pp.17-24.

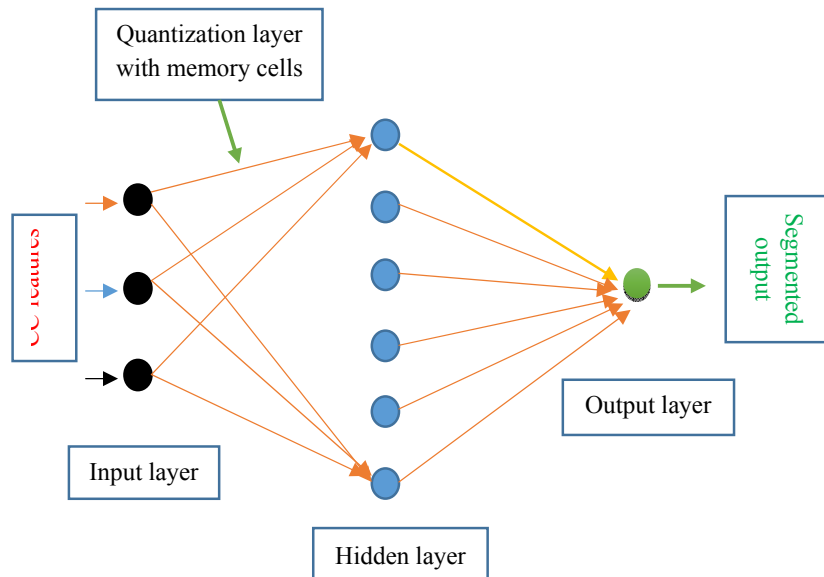


Fig.1 CMAC for Segmentation of Lung slice image

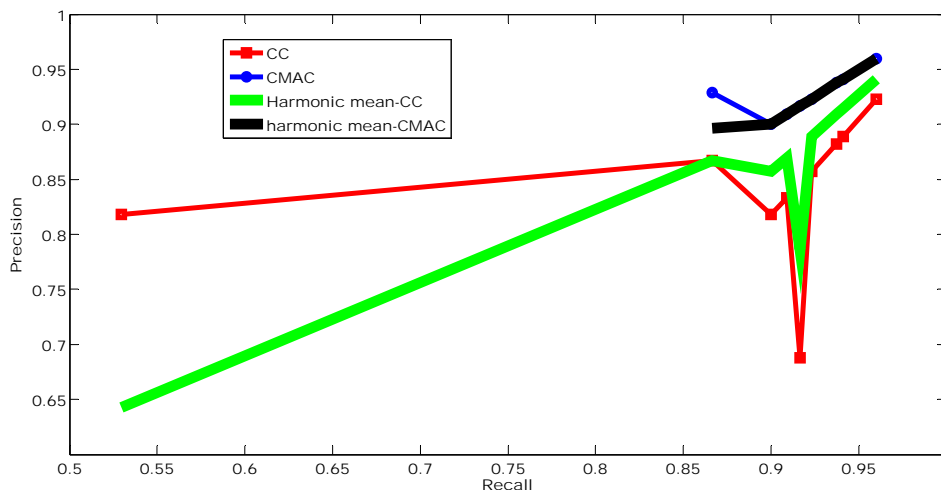
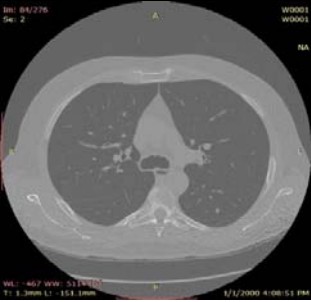
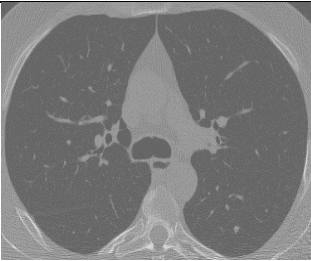
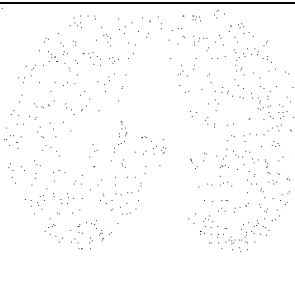
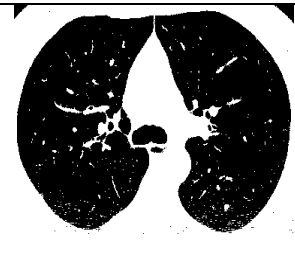
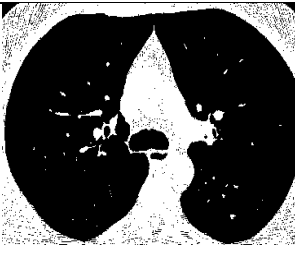


Fig.2 Precision-recall curve

Table 4.1 Segmented results of block size 3x3 for threshold 20 to 160

Node in the hidden layer	Image	Description
		Original image
		Cropped original
2		Poor segmentation

4		Lungs are visible with little segmentation.
6		Lungs are demarcated from the background
8		Lungs are clearly visible with least noise inside them



Dr. S. Purushothaman completed his PhD from Indian Institute of Technology Madras, India in 1995. He has 168 publications to his credit. He has 23 years of teaching experience. Presently he is working as Associate Professor in Institute of Technology, Haramaya University, Ethiopia.



Benita K. J. Veronica is doing research in Mother Teresa Women's University, Kodaikanal, India. Her areas of research interest is Image Processing and Neural Network. Having 4 years of teaching experience.



Dr. P. Rajeswari completed her PhD in Mother Teresa Women's University, Kodaikanal, India in 2014. She has 32 publications to her credit. Her areas of research interest is Intelligent Computing.

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore, Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGILIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Dr. Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Dr Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Dr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Dr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Dr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Dr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Dr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Dr. P. Vasant, University Technology Petronas, Malaysia
Dr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Dr. Praveen Ranjan Srivastava, BITS PILANI, India
Dr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Dr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Dr. Tirthankar Gayen, IIT Kharagpur, India
Dr. Huei-Ru Tseng, National Chiao Tung University, Taiwan
Prof. Ning Xu, Wuhan University of Technology, China
Dr Mohammed Salem Binwahlan, Hadhramout University of Science and Technology, Yemen
& Universiti Teknologi Malaysia, Malaysia.
Dr. Aruna Ranganath, Bhoj Reddy Engineering College for Women, India
Dr. Hafeezullah Amin, Institute of Information Technology, KUST, Kohat, Pakistan

Prof. Syed S. Rizvi, University of Bridgeport, USA
Dr. Shahbaz Pervez Chattha, University of Engineering and Technology Taxila, Pakistan
Dr. Shishir Kumar, Jaypee University of Information Technology, Wakanaghat (HP), India
Dr. Shahid Mumtaz, Portugal Telecommunication, Instituto de Telecomunicações (IT) , Aveiro, Portugal
Dr. Rajesh K Shukla, Corporate Institute of Science & Technology Bhopal M P
Dr. Poonam Garg, Institute of Management Technology, India
Dr. S. Mehta, Inha University, Korea
Dr. Dilip Kumar S.M, Bangalore University, Bangalore
Prof. Malik Sikander Hayat Khiyal, Fatima Jinnah Women University, Rawalpindi, Pakistan
Dr. Virendra Gomase , Department of Bioinformatics, Padmashree Dr. D.Y. Patil University
Dr. Irraivan Elamvazuthi, University Technology PETRONAS, Malaysia
Dr. Saqib Saeed, University of Siegen, Germany
Dr. Pavan Kumar Gorakavi, IPMA-USA [YC]
Dr. Ahmed Nabih Zaki Rashed, Menoufia University, Egypt
Prof. Shishir K. Shandilya, Rukmani Devi Institute of Science & Technology, India
Dr. J. Komala Lakshmi, SNR Sons College, Computer Science, India
Dr. Muhammad Sohail, KUST, Pakistan
Dr. Manjaiah D.H, Mangalore University, India
Dr. S Santhosh Baboo, D.G.Vaishnav College, Chennai, India
Prof. Dr. Mokhtar Beldjehem, Sainte-Anne University, Halifax, NS, Canada
Dr. Deepak Laxmi Narasimha, University of Malaya, Malaysia
Prof. Dr. Arunkumar Thangavelu, Vellore Institute Of Technology, India
Dr. M. Azath, Anna University, India
Dr. Md. Rabiul Islam, Rajshahi University of Engineering & Technology (RUET), Bangladesh
Dr. Aos Alaa Zaidan Ansaef, Multimedia University, Malaysia
Dr Suresh Jain, Devi Ahilya University, Indore (MP) India,
Dr. Mohammed M. Kadhum, Universiti Utara Malaysia
Dr. Hanumanthappa. J. University of Mysore, India
Dr. Syed Ishtiaque Ahmed, Bangladesh University of Engineering and Technology (BUET)
Dr Akinola Solomon Olalekan, University of Ibadan, Ibadan, Nigeria
Dr. Santosh K. Pandey, The Institute of Chartered Accountants of India
Dr. P. Vasant, Power Control Optimization, Malaysia
Dr. Petr Ivankov, Automatika - S, Russian Federation
Dr. Utkarsh Seetha, Data Infosys Limited, India
Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal
Dr. (Mrs) Padmavathi Ganapathi, Avinashilingam University for Women, Coimbatore
Assist. Prof. A. Neela madheswari, Anna university, India
Prof. Ganesan Ramachandra Rao, PSG College of Arts and Science, India
Mr. Kamanashis Biswas, Daffodil International University, Bangladesh
Dr. Atul Gonsai, Saurashtra University, Gujarat, India
Mr. Angkoon Phinyomark, Prince of Songkla University, Thailand
Mrs. G. Nalini Priya, Anna University, Chennai
Dr. P. Subashini, Avinashilingam University for Women, India
Assoc. Prof. Vijay Kumar Chakka, Dhirubhai Ambani IICT, Gandhinagar ,Gujarat
Mr Jitendra Agrawal, : Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal
Mr. Vishal Goyal, Department of Computer Science, Punjabi University, India
Dr. R. Baskaran, Department of Computer Science and Engineering, Anna University, Chennai

Assist. Prof. Kanwalvir Singh Dhindsa, B.B.S.B.Engg.College, Fatehgarh Sahib (Punjab), India
Dr. Jamal Ahmad Dargham, School of Engineering and Information Technology, Universiti Malaysia Sabah
Mr. Nitin Bhatia, DAV College, India
Dr. Dhavachelvan Ponnurangam, Pondicherry Central University, India
Dr. Mohd Faizal Abdollah, University of Technical Malaysia, Malaysia
Assist. Prof. Sonal Chawla, Panjab University, India
Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India
Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia
Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia
Professor Dr. Sabu M. Thampi, .B.S Institute of Technology for Women, Kerala University, India
Mr. Noor Muhammed Nayeem, Université Lumière Lyon 2, 69007 Lyon, France
Dr. Himanshu Aggarwal, Department of Computer Engineering, Punjabi University, India
Prof R. Naidoo, Dept of Mathematics/Center for Advanced Computer Modelling, Durban University of Technology,
Durban, South Africa
Prof. Mydhili K Nair, Visweswaraiah Technological University, Bangalore, India
M. Prabu, Adhiyamaan College of Engineering/Anna University, India
Mr. Swakkhar Shatabda, United International University, Bangladesh
Dr. Abdur Rashid Khan, ICIT, Gomal University, Dera Ismail Khan, Pakistan
Mr. H. Abdul Shabeer, I-Nautix Technologies, Chennai, India
Dr. M. Aramudhan, Perunthalaivar Kamarajar Institute of Engineering and Technology, India
Dr. M. P. Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), India
Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran
Mr. Zeashan Hameed Khan, Université de Grenoble, France
Prof. Anil K Ahlawat, Ajay Kumar Garg Engineering College, Ghaziabad, UP Technical University, Lucknow
Mr. Longe Olumide Babatope, University Of Ibadan, Nigeria
Associate Prof. Raman Maini, University College of Engineering, Punjabi University, India
Dr. Maslin Masrom, University Technology Malaysia, Malaysia
Sudipta Chattopadhyay, Jadavpur University, Kolkata, India
Dr. Dang Tuan NGUYEN, University of Information Technology, Vietnam National University - Ho Chi Minh City
Dr. Mary Lourde R., BITS-PILANI Dubai , UAE
Dr. Abdul Aziz, University of Central Punjab, Pakistan
Mr. Karan Singh, Gautam Budtha University, India
Mr. Avinash Pokhriyal, Uttar Pradesh Technical University, Lucknow, India
Associate Prof Dr Zuraini Ismail, University Technology Malaysia, Malaysia
Assistant Prof. Yasser M. Alginahi, Taibah University, Madinah Munawwarah, KSA
Mr. Dakshina Ranjan Kisku, West Bengal University of Technology, India
Mr. Raman Kumar, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India
Associate Prof. Samir B. Patel, Institute of Technology, Nirma University, India
Dr. M.Munir Ahamed Rabbani, B. S. Abdur Rahman University, India
Asst. Prof. Koushik Majumder, West Bengal University of Technology, India
Dr. Alex Pappachen James, Queensland Micro-nanotechnology center, Griffith University, Australia
Assistant Prof. S. Hariharan, B.S. Abdur Rahman University, India
Asst Prof. Jasmine. K. S, R.V.College of Engineering, India
Mr Naushad Ali Mamode Khan, Ministry of Education and Human Resources, Mauritius
Prof. Mahesh Goyani, G H Patel Collge of Engg. & Tech, V.V.N, Anand, Gujarat, India
Dr. Mana Mohammed, University of Tlemcen, Algeria
Prof. Jatinder Singh, Universal Institution of Engg. & Tech. CHD, India

Mrs. M. Anandhavalli Gauthaman, Sikkim Manipal Institute of Technology, Majitar, East Sikkim
Dr. Bin Guo, Institute Telecom SudParis, France
Mrs. Maleika Mehr Nigar Mohamed Heenaye-Mamode Khan, University of Mauritius
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Mr. V. Bala Dhandayuthapani, Mekelle University, Ethiopia
Dr. Irfan Syamsuddin, State Polytechnic of Ujung Pandang, Indonesia
Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius
Mr. Ravi Chandiran, Zagro Singapore Pte Ltd. Singapore
Mr. Miliindkumar V. Sarode, Jawaharlal Darda Institute of Engineering and Technology, India
Dr. Shamimul Qamar, KSJ Institute of Engineering & Technology, India
Dr. C. Arun, Anna University, India
Assist. Prof. M.N.Birje, Basaveshwar Engineering College, India
Prof. Hamid Reza Naji, Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran
Assist. Prof. Debasis Giri, Department of Computer Science and Engineering, Haldia Institute of Technology
Subhabrata Barman, Haldia Institute of Technology, West Bengal
Mr. M. I. Lali, COMSATS Institute of Information Technology, Islamabad, Pakistan
Dr. Feroz Khan, Central Institute of Medicinal and Aromatic Plants, Lucknow, India
Mr. R. Nagendran, Institute of Technology, Coimbatore, Tamilnadu, India
Mr. Amnach Khawne, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand
Dr. P. Chakrabarti, Sir Padampat Singhania University, Udaipur, India
Mr. Nafiz Imtiaz Bin Hamid, Islamic University of Technology (IUT), Bangladesh.
Shahab-A. Shamshirband, Islamic Azad University, Chalous, Iran
Prof. B. Priestly Shan, Anna Univeristy, Tamilnadu, India
Venkatramreddy Velma, Dept. of Bioinformatics, University of Mississippi Medical Center, Jackson MS USA
Akshi Kumar, Dept. of Computer Engineering, Delhi Technological University, India
Dr. Umesh Kumar Singh, Vikram University, Ujjain, India
Mr. Serguei A. Mokhov, Concordia University, Canada
Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia
Dr. Awadhesh Kumar Sharma, Madan Mohan Malviya Engineering College, India
Mr. Syed R. Rizvi, Analytical Services & Materials, Inc., USA
Dr. S. Karthik, SNS College of Technology, India
Mr. Syed Qasim Bukhari, CIMET (Universidad de Granada), Spain
Mr. A.D.Potgantwar, Pune University, India
Dr. Himanshu Aggarwal, Punjabi University, India
Mr. Rajesh Ramachandran, Naipunya Institute of Management and Information Technology, India
Dr. K.L. Shunmuganathan, R.M.K Engg College , Kavaraipettai ,Chennai
Dr. Prasant Kumar Pattnaik, KIST, India.
Dr. Ch. Aswani Kumar, VIT University, India
Mr. Ijaz Ali Shoukat, King Saud University, Riyadh KSA
Mr. Arun Kumar, Sir Padam Pat Singhania University, Udaipur, Rajasthan
Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia
Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA
Mr. Mohd Zaki Bin Mas'ud, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
Prof. Dr. R. Geetharamani, Dept. of Computer Science and Eng., Rajalakshmi Engineering College, India
Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India
Dr. S. Abdul Khader Jilani, University of Tabuk, Tabuk, Saudi Arabia
Mr. Syed Jamal Haider Zaidi, Bahria University, Pakistan

Dr. N. Devarajan, Government College of Technology, Coimbatore, Tamilnadu, INDIA
Mr. R. Jagadeesh Kannan, RMK Engineering College, India
Mr. Deo Prakash, Shri Mata Vaishno Devi University, India
Mr. Mohammad Abu Naser, Dept. of EEE, IUT, Gazipur, Bangladesh
Assist. Prof. Prasun Ghosal, Bengal Engineering and Science University, India
Mr. Md. Golam Kaosar, School of Engineering and Science, Victoria University, Melbourne City, Australia
Mr. R. Mohammad Shafi, Madanapalle Institute of Technology & Science, India
Dr. F. Sagayaraj Francis, Pondicherry Engineering College, India
Dr. Ajay Goel, HIET, Kaithal, India
Mr. Nayak Sunil Kashibarao, Bahirji Smarak Mahavidyalaya, India
Mr. Suhas J Manangi, Microsoft India
Dr. Kalyankar N. V., Yeshwant Mahavidyalaya, Nanded, India
Dr. K.D. Verma, S.V. College of Post graduate studies & Research, India
Dr. Amjad Rehman, University Technology Malaysia, Malaysia
Mr. Rachit Garg, L K College, Jalandhar, Punjab
Mr. J. William, M.A.M college of Engineering, Trichy, Tamilnadu, India
Prof. Jue-Sam Chou, Nanhua University, College of Science and Technology, Taiwan
Dr. Thorat S.B., Institute of Technology and Management, India
Mr. Ajay Prasad, Sir Padampat Singhanian University, Udaipur, India
Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology & Science, India
Mr. Syed Rafiul Hussain, Ahsanullah University of Science and Technology, Bangladesh
Mrs Fazeela Tunnisa, Najran University, Kingdom of Saudi Arabia
Mrs Kavita Taneja, Maharishi Markandeshwar University, Haryana, India
Mr. Maniyar Shiraz Ahmed, Najran University, Najran, KSA
Mr. Anand Kumar, AMC Engineering College, Bangalore
Dr. Rakesh Chandra Gangwar, Beant College of Engg. & Tech., Gurdaspur (Punjab) India
Dr. V V Rama Prasad, Sree Vidyanikethan Engineering College, India
Assist. Prof. Neetesh Kumar Gupta, Technocrats Institute of Technology, Bhopal (M.P.), India
Mr. Ashish Seth, Uttar Pradesh Technical University, Lucknow, UP India
Dr. V V S S S Balam, Sreenidhi Institute of Science and Technology, India
Mr Rahul Bhatia, Lingaya's Institute of Management and Technology, India
Prof. Niranjana Reddy, P, KITS, Warangal, India
Prof. Rakesh. Lingappa, Vijetha Institute of Technology, Bangalore, India
Dr. Mohammed Ali Hussain, Nimra College of Engineering & Technology, Vijayawada, A.P., India
Dr. A. Srinivasan, MNM Jain Engineering College, Rajiv Gandhi Salai, Thorapakkam, Chennai
Mr. Rakesh Kumar, M.M. University, Mullana, Ambala, India
Dr. Lena Khaled, Zarqa Private University, Aman, Jordan
Ms. Supriya Kapoor, Patni/Lingaya's Institute of Management and Tech., India
Dr. Tossapon Boongoen, Aberystwyth University, UK
Dr. Bilal Alatas, Firat University, Turkey
Assist. Prof. Jyoti Praaksh Singh, Academy of Technology, India
Dr. Ritu Soni, GNG College, India
Dr. Mahendra Kumar, Sagar Institute of Research & Technology, Bhopal, India.
Dr. Binod Kumar, Lakshmi Narayan College of Tech. (LNCT) Bhopal India
Dr. Muzhir Shaban Al-Ani, Amman Arab University Amman – Jordan
Dr. T.C. Manjunath, ATRIA Institute of Tech, India
Mr. Muhammad Zakarya, COMSATS Institute of Information Technology (CIIT), Pakistan

Assist. Prof. Harmunish Taneja, M. M. University, India
Dr. Chitra Dhawale , SICSR, Model Colony, Pune, India
Mrs Sankari Muthukaruppan, Nehru Institute of Engineering and Technology, Anna University, India
Mr. Aaqif Afzaal Abbasi, National University Of Sciences And Technology, Islamabad
Prof. Ashutosh Kumar Dubey, Trinity Institute of Technology and Research Bhopal, India
Mr. G. Appasami, Dr. Pauls Engineering College, India
Mr. M Yasin, National University of Science and Tech, karachi (NUST), Pakistan
Mr. Yaser Miaji, University Utara Malaysia, Malaysia
Mr. Shah Ahsanul Haque, International Islamic University Chittagong (IIUC), Bangladesh
Prof. (Dr) Syed Abdul Sattar, Royal Institute of Technology & Science, India
Dr. S. Sasikumar, Roever Engineering College
Assist. Prof. Monit Kapoor, Maharishi Markandeshwar University, India
Mr. Nwaocha Vivian O, National Open University of Nigeria
Dr. M. S. Vijaya, GR Govindarajulu School of Applied Computer Technology, India
Assist. Prof. Chakresh Kumar, Manav Rachna International University, India
Mr. Kunal Chadha , R&D Software Engineer, Gemalto, Singapore
Mr. Mueen Uddin, Universiti Teknologi Malaysia, UTM , Malaysia
Dr. Dhuha Basheer abdullah, Mosul university, Iraq
Mr. S. Audithan, Annamalai University, India
Prof. Vijay K Chaudhari, Technocrats Institute of Technology , India
Associate Prof. Mohd Ilyas Khan, Technocrats Institute of Technology , India
Dr. Vu Thanh Nguyen, University of Information Technology, HoChiMinh City, VietNam
Assist. Prof. Anand Sharma, MITS, Lakshmarangarh, Sikar, Rajasthan, India
Prof. T V Narayana Rao, HITAM Engineering college, Hyderabad
Mr. Deepak Gour, Sir Padampat Singhania University, India
Assist. Prof. Amutharaj Joyson, Kalasalingam University, India
Mr. Ali Balador, Islamic Azad University, Iran
Mr. Mohit Jain, Maharaja Surajmal Institute of Technology, India
Mr. Dilip Kumar Sharma, GLA Institute of Technology & Management, India
Dr. Debojyoti Mitra, Sir padampat Singhania University, India
Dr. Ali Dehghantanha, Asia-Pacific University College of Technology and Innovation, Malaysia
Mr. Zhao Zhang, City University of Hong Kong, China
Prof. S.P. Setty, A.U. College of Engineering, India
Prof. Patel Rakeshkumar Kantilal, Sankalchand Patel College of Engineering, India
Mr. Biswajit Bhowmik, Bengal College of Engineering & Technology, India
Mr. Manoj Gupta, Apex Institute of Engineering & Technology, India
Assist. Prof. Ajay Sharma, Raj Kumar Goel Institute Of Technology, India
Assist. Prof. Ramveer Singh, Raj Kumar Goel Institute of Technology, India
Dr. Hanan Elazhary, Electronics Research Institute, Egypt
Dr. Hosam I. Faiq, USM, Malaysia
Prof. Dipti D. Patil, MAEER's MIT College of Engg. & Tech, Pune, India
Assist. Prof. Devendra Chack, BCT Kumaon engineering College Dwarahat Almora, India
Prof. Manpreet Singh, M. M. Engg. College, M. M. University, India
Assist. Prof. M. Sadiq ali Khan, University of Karachi, Pakistan
Mr. Prasad S. Halgaonkar, MIT - College of Engineering, Pune, India
Dr. Imran Ghani, Universiti Teknologi Malaysia, Malaysia
Prof. Varun Kumar Kakar, Kumaon Engineering College, Dwarahat, India

Assist. Prof. Nisheeth Joshi, Apaji Institute, Banasthali University, Rajasthan, India
Associate Prof. Kunwar S. Vaisla, VCT Kumaon Engineering College, India
Prof Anupam Choudhary, Bhilai School Of Engg.,Bhilai (C.G.),India
Mr. Divya Prakash Shrivastava, Al Jabal Al garbi University, Zawya, Libya
Associate Prof. Dr. V. Radha, Avinashilingam Deemed university for women, Coimbatore.
Dr. Kasarapu Ramani, JNT University, Anantapur, India
Dr. Anuraag Awasthi, Jayoti Vidyapeeth Womens University, India
Dr. C G Ravichandran, R V S College of Engineering and Technology, India
Dr. Mohamed A. Deriche, King Fahd University of Petroleum and Minerals, Saudi Arabia
Mr. Abbas Karimi, Universiti Putra Malaysia, Malaysia
Mr. Amit Kumar, Jaypee University of Engg. and Tech., India
Dr. Nikolai Stoianov, Defense Institute, Bulgaria
Assist. Prof. S. Ranichandra, KSR College of Arts and Science, Tiruchencode
Mr. T.K.P. Rajagopal, Diamond Horse International Pvt Ltd, India
Dr. Md. Ekramul Hamid, Rajshahi University, Bangladesh
Mr. Hemanta Kumar Kalita , TATA Consultancy Services (TCS), India
Dr. Messaouda Azzouzi, Ziane Achour University of Djelfa, Algeria
Prof. (Dr.) Juan Jose Martinez Castillo, "Gran Mariscal de Ayacucho" University and Acantelys research Group, Venezuela
Dr. Jatinderkumar R. Saini, Narmada College of Computer Application, India
Dr. Babak Bashari Rad, University Technology of Malaysia, Malaysia
Dr. Nighat Mir, Effat University, Saudi Arabia
Prof. (Dr.) G.M.Nasira, Sasurie College of Engineering, India
Mr. Varun Mittal, Gemalto Pte Ltd, Singapore
Assist. Prof. Mrs P. Banumathi, Kathir College Of Engineering, Coimbatore
Assist. Prof. Quan Yuan, University of Wisconsin-Stevens Point, US
Dr. Pranam Paul, Narula Institute of Technology, Agarpara, West Bengal, India
Assist. Prof. J. Ramkumar, V.L.B Janakiammal college of Arts & Science, India
Mr. P. Sivakumar, Anna university, Chennai, India
Mr. Md. Humayun Kabir Biswas, King Khalid University, Kingdom of Saudi Arabia
Mr. Mayank Singh, J.P. Institute of Engg & Technology, Meerut, India
HJ. Kamaruzaman Jusoff, Universiti Putra Malaysia
Mr. Nikhil Patrick Lobo, CADES, India
Dr. Amit Wason, Rayat-Bahra Institute of Engineering & Boi-Technology, India
Dr. Rajesh Shrivastava, Govt. Benazir Science & Commerce College, Bhopal, India
Assist. Prof. Vishal Bharti, DCE, Gurgaon
Mrs. Sunita Bansal, Birla Institute of Technology & Science, India
Dr. R. Sudhakar, Dr.Mahalingam college of Engineering and Technology, India
Dr. Amit Kumar Garg, Shri Mata Vaishno Devi University, Katra(J&K), India
Assist. Prof. Raj Gaurang Tiwari, AZAD Institute of Engineering and Technology, India
Mr. Hamed Taherdoost, Tehran, Iran
Mr. Amin Daneshmand Malayeri, YRC, IAU, Malayer Branch, Iran
Mr. Shantanu Pal, University of Calcutta, India
Dr. Terry H. Walcott, E-Promag Consultancy Group, United Kingdom
Dr. Ezekiel U OKIKE, University of Ibadan, Nigeria
Mr. P. Mahalingam, Caledonian College of Engineering, Oman
Dr. Mahmoud M. A. Abd Ellatif, Mansoura University, Egypt

Prof. Kunwar S. Vaisla, BCT Kumaon Engineering College, India
Prof. Mahesh H. Panchal, Kalol Institute of Technology & Research Centre, India
Mr. Muhammad Asad, Technical University of Munich, Germany
Mr. AliReza Shams Shafigh, Azad Islamic university, Iran
Prof. S. V. Nagaraj, RMK Engineering College, India
Mr. Ashikali M Hasan, Senior Researcher, CelNet security, India
Dr. Adnan Shahid Khan, University Technology Malaysia, Malaysia
Mr. Prakash Gajanan Burade, Nagpur University/ITM college of engg, Nagpur, India
Dr. Jagdish B.Helonde, Nagpur University/ITM college of engg, Nagpur, India
Professor, Doctor BOUHORMA Mohammed, Univertsity Abdelmalek Essaadi, Morocco
Mr. K. Thirumalaivasan, Pondicherry Engg. College, India
Mr. Umbarkar Anantkumar Janardan, Walchand College of Engineering, India
Mr. Ashish Chaurasia, Gyan Ganga Institute of Technology & Sciences, India
Mr. Sunil Taneja, Kurukshetra University, India
Mr. Fauzi Adi Rafrastara, Dian Nuswantoro University, Indonesia
Dr. Yaduvir Singh, Thapar University, India
Dr. Ioannis V. Koskosas, University of Western Macedonia, Greece
Dr. Vasantha Kalyani David, Avinashilingam University for women, Coimbatore
Dr. Ahmed Mansour Manasrah, Universiti Sains Malaysia, Malaysia
Miss. Nazanin Sadat Kazazi, University Technology Malaysia, Malaysia
Mr. Saeed Rasouli Heikalabad, Islamic Azad University - Tabriz Branch, Iran
Assoc. Prof. Dharendra Mishra, SVKM's NMIMS University, India
Prof. Shapoor Zarei, UAE Inventors Association, UAE
Prof. B.Raja Sarath Kumar, Lenora College of Engineering, India
Dr. Bashir Alam, Jamia millia Islamia, Delhi, India
Prof. Anant J Umbarkar, Walchand College of Engg., India
Assist. Prof. B. Bharathi, Sathyabama University, India
Dr. Fokrul Alom Mazarbhuiya, King Khalid University, Saudi Arabia
Prof. T.S.Jeyali Laseeth, Anna University of Technology, Tirunelveli, India
Dr. M. Balraju, Jawahar Lal Nehru Technological University Hyderabad, India
Dr. Vijayalakshmi M. N., R.V.College of Engineering, Bangalore
Prof. Walid Moudani, Lebanese University, Lebanon
Dr. Saurabh Pal, VBS Purvanchal University, Jaunpur, India
Associate Prof. Suneet Chaudhary, Dehradun Institute of Technology, India
Associate Prof. Dr. Manuj Darbari, BBD University, India
Ms. Prema Selvaraj, K.S.R College of Arts and Science, India
Assist. Prof. Ms.S.Sasikala, KSR College of Arts & Science, India
Mr. Sukhvinder Singh Deora, NC Institute of Computer Sciences, India
Dr. Abhay Bansal, Amity School of Engineering & Technology, India
Ms. Sumita Mishra, Amity School of Engineering and Technology, India
Professor S. Viswanadha Raju, JNT University Hyderabad, India
Mr. Asghar Shahrzad Khashandarag, Islamic Azad University Tabriz Branch, India
Mr. Manoj Sharma, Panipat Institute of Engg. & Technology, India
Mr. Shakeel Ahmed, King Faisal University, Saudi Arabia
Dr. Mohamed Ali Mahjoub, Institute of Engineer of Monastir, Tunisia
Mr. Adri Jovin J.J., SriGuru Institute of Technology, India
Dr. Sukumar Senthikumar, Universiti Sains Malaysia, Malaysia

Mr. Rakesh Bharati, Dehradun Institute of Technology Dehradun, India
Mr. Shervan Fekri Ershad, Shiraz International University, Iran
Mr. Md. Safiqul Islam, Daffodil International University, Bangladesh
Mr. Mahmudul Hasan, Daffodil International University, Bangladesh
Prof. Mandakini Tayade, UIT, RGTU, Bhopal, India
Ms. Sarla More, UIT, RGTU, Bhopal, India
Mr. Tushar Hrishikesh Jaware, R.C. Patel Institute of Technology, Shirpur, India
Ms. C. Divya, Dr G R Damodaran College of Science, Coimbatore, India
Mr. Fahimuddin Shaik, Annamacharya Institute of Technology & Sciences, India
Dr. M. N. Giri Prasad, JNTUCE, Pulivendula, A.P., India
Assist. Prof. Chintan M Bhatt, Charotar University of Science And Technology, India
Prof. Sahista Machchhar, Marwadi Education Foundation's Group of institutions, India
Assist. Prof. Navnish Goel, S. D. College Of Engineering & Technology, India
Mr. Khaja Kamaluddin, Sirt University, Sirt, Libya
Mr. Mohammad Zaidul Karim, Daffodil International, Bangladesh
Mr. M. Vijayakumar, KSR College of Engineering, Tiruchengode, India
Mr. S. A. Ahsan Rajon, Khulna University, Bangladesh
Dr. Muhammad Mohsin Nazir, LCW University Lahore, Pakistan
Mr. Mohammad Asadul Hoque, University of Alabama, USA
Mr. P.V.Sarathchand, Indur Institute of Engineering and Technology, India
Mr. Durgesh Samadhiya, Chung Hua University, Taiwan
Dr Venu Kuthadi, University of Johannesburg, Johannesburg, RSA
Dr. (Er) Jasvir Singh, Guru Nanak Dev University, Amritsar, Punjab, India
Mr. Jasmin Cosic, Min. of the Interior of Una-sana canton, B&H, Bosnia and Herzegovina
Dr S. Rajalakshmi, Botho College, South Africa
Dr. Mohamed Sarrab, De Montfort University, UK
Mr. Basappa B. Kodada, Canara Engineering College, India
Assist. Prof. K. Ramana, Annamacharya Institute of Technology and Sciences, India
Dr. Ashu Gupta, Apeejay Institute of Management, Jalandhar, India
Assist. Prof. Shaik Rasool, Shadan College of Engineering & Technology, India
Assist. Prof. K. Suresh, Annamacharya Institute of Tech & Sci. Rajampet, AP, India
Dr . G. Singaravel, K.S.R. College of Engineering, India
Dr B. G. Geetha, K.S.R. College of Engineering, India
Assist. Prof. Kavita Choudhary, ITM University, Gurgaon
Dr. Mehrdad Jalali, Azad University, Mashhad, Iran
Megha Goel, Shamli Institute of Engineering and Technology, Shamli, India
Mr. Chi-Hua Chen, Institute of Information Management, National Chiao-Tung University, Taiwan (R.O.C.)
Assoc. Prof. A. Rajendran, RVS College of Engineering and Technology, India
Assist. Prof. S. Jaganathan, RVS College of Engineering and Technology, India
Assoc. Prof. (Dr.) A S N Chakravarthy, JNTUK University College of Engineering Vizianagaram (State University)
Assist. Prof. Deepshikha Patel, Technocrat Institute of Technology, India
Assist. Prof. Maram Balajee, GMRIT, India
Assist. Prof. Monika Bhatnagar, TIT, India
Prof. Gaurang Panchal, Charotar University of Science & Technology, India
Prof. Anand K. Tripathi, Computer Society of India
Prof. Jyoti Chaudhary, High Performance Computing Research Lab, India
Assist. Prof. Supriya Raheja, ITM University, India

Dr. Pankaj Gupta, Microsoft Corporation, U.S.A.
Assist. Prof. Panchamukesh Chandaka, Hyderabad Institute of Tech. & Management, India
Prof. Mohan H.S, SJB Institute Of Technology, India
Mr. Hossein Malekinezhad, Islamic Azad University, Iran
Mr. Zatin Gupta, Universti Malaysia, Malaysia
Assist. Prof. Amit Chauhan, Phonics Group of Institutions, India
Assist. Prof. Ajal A. J., METS School Of Engineering, India
Mrs. Omowunmi Omobola Adeyemo, University of Ibadan, Nigeria
Dr. Bharat Bhushan Agarwal, I.F.T.M. University, India
Md. Nazrul Islam, University of Western Ontario, Canada
Tushar Kanti, L.N.C.T, Bhopal, India
Er. Aumreesh Kumar Saxena, SIRTs College Bhopal, India
Mr. Mohammad Monirul Islam, Daffodil International University, Bangladesh
Dr. Kashif Nisar, University Utara Malaysia, Malaysia
Dr. Wei Zheng, Rutgers Univ/ A10 Networks, USA
Associate Prof. Rituraj Jain, Vyas Institute of Engg & Tech, Jodhpur – Rajasthan
Assist. Prof. Apoorvi Sood, I.T.M. University, India
Dr. Kayhan Zrar Ghafoor, University Technology Malaysia, Malaysia
Mr. Swapnil Sonar, Truba Institute College of Engineering & Technology, Indore, India
Ms. Yogita Gigras, I.T.M. University, India
Associate Prof. Neelima Sadineni, Pydha Engineering College, India Pydha Engineering College
Assist. Prof. K. Deepika Rani, HITAM, Hyderabad
Ms. Shikha Maheshwari, Jaipur Engineering College & Research Centre, India
Prof. Dr V S Giridhar Akula, Avanthi's Scientific Tech. & Research Academy, Hyderabad
Prof. Dr.S.Saravanan, Muthayammal Engineering College, India
Mr. Mehdi Golsorkhatabar Amiri, Islamic Azad University, Iran
Prof. Amit Sadanand Savyanavar, MITCOE, Pune, India
Assist. Prof. P.Oliver Jayaprakash, Anna University, Chennai
Assist. Prof. Ms. Sujata, ITM University, Gurgaon, India
Dr. Asoke Nath, St. Xavier's College, India
Mr. Masoud Rafiqhi, Islamic Azad University, Iran
Assist. Prof. RamBabu Pemula, NIMRA College of Engineering & Technology, India
Assist. Prof. Ms Rita Chhikara, ITM University, Gurgaon, India
Mr. Sandeep Maan, Government Post Graduate College, India
Prof. Dr. S. Muralidharan, Mepco Schlenk Engineering College, India
Associate Prof. T.V.Sai Krishna, QIS College of Engineering and Technology, India
Mr. R. Balu, Bharathiar University, Coimbatore, India
Assist. Prof. Shekhar. R, Dr.SM College of Engineering, India
Prof. P. Senthilkumar, Vivekanandha Institue of Engineering and Techology for Woman, India
Mr. M. Kamarajan, PSNA College of Engineering & Technology, India
Dr. Angajala Srinivasa Rao, Jawaharlal Nehru Technical University, India
Assist. Prof. C. Venkatesh, A.I.T.S, Rajampet, India
Mr. Afshin Rezakhani Roozbahani, Ayatollah Boroujerdi University, Iran
Mr. Laxmi chand, SCTL, Noida, India
Dr. Dr. Abdul Hannan, Vivekanand College, Aurangabad
Prof. Mahesh Panchal, KITRC, Gujarat
Dr. A. Subramani, K.S.R. College of Engineering, Tiruchengode

Assist. Prof. Prakash M, Rajalakshmi Engineering College, Chennai, India
Assist. Prof. Akhilesh K Sharma, Sir Padampat Singhanian University, India
Ms. Varsha Sahni, Guru Nanak Dev Engineering College, Ludhiana, India
Associate Prof. Trilochan Rout, NM Institute of Engineering and Technology, India
Mr. Srikanta Kumar Mohapatra, NMIET, Orissa, India
Mr. Waqas Haider Bangyal, Iqra University Islamabad, Pakistan
Dr. S. Vijayaragavan, Christ College of Engineering and Technology, Pondicherry, India
Prof. Elboukhari Mohamed, University Mohammed First, Oujda, Morocco
Dr. Muhammad Asif Khan, King Faisal University, Saudi Arabia
Dr. Nagy Ramadan Darwish Omran, Cairo University, Egypt.
Assistant Prof. Anand Nayyar, KCL Institute of Management and Technology, India
Mr. G. Premsankar, Ericsson, India
Assist. Prof. T. Hemalatha, VELS University, India
Prof. Tejaswini Apte, University of Pune, India
Dr. Edmund Ng Giap Weng, Universiti Malaysia Sarawak, Malaysia
Mr. Mahdi Nouri, Iran University of Science and Technology, Iran
Associate Prof. S. Asif Hussain, Annamacharya Institute of technology & Sciences, India
Mrs. Kavita Pabreja, Maharaja Surajmal Institute (an affiliate of GGSIP University), India
Mr. Vorugunti Chandra Sekhar, DA-IICT, India
Mr. Muhammad Najmi Ahmad Zabidi, Universiti Teknologi Malaysia, Malaysia
Dr. Aderemi A. Atayero, Covenant University, Nigeria
Assist. Prof. Osama Sohaib, Balochistan University of Information Technology, Pakistan
Assist. Prof. K. Suresh, Annamacharya Institute of Technology and Sciences, India
Mr. Hassen Mohammed Abdullaah Alsafi, International Islamic University Malaysia (IIUM) Malaysia
Mr. Robail Yasrab, Virtual University of Pakistan, Pakistan
Mr. R. Balu, Bharathiar University, Coimbatore, India
Prof. Anand Nayyar, KCL Institute of Management and Technology, Jalandhar
Assoc. Prof. Vivek S Deshpande, MIT College of Engineering, India
Prof. K. Saravanan, Anna university Coimbatore, India
Dr. Ravendra Singh, MJP Rohilkhand University, Bareilly, India
Mr. V. Mathivanan, IBRA College of Technology, Sultanate of OMAN
Assoc. Prof. S. Asif Hussain, AITS, India
Assist. Prof. C. Venkatesh, AITS, India
Mr. Sami Ulhaq, SZABIST Islamabad, Pakistan
Dr. B. Justus Rabi, Institute of Science & Technology, India
Mr. Anuj Kumar Yadav, Dehradun Institute of technology, India
Mr. Alejandro Mosquera, University of Alicante, Spain
Assist. Prof. Arjun Singh, Sir Padampat Singhanian University (SPSU), Udaipur, India
Dr. Smriti Agrawal, JB Institute of Engineering and Technology, Hyderabad
Assist. Prof. Swathi Sambangi, Visakha Institute of Engineering and Technology, India
Ms. Prabhjot Kaur, Guru Gobind Singh Indraprastha University, India
Mrs. Samaher AL-Hothali, Yanbu University College, Saudi Arabia
Prof. Rajneeshkaur Bedi, MIT College of Engineering, Pune, India
Mr. Hassen Mohammed Abdullaah Alsafi, International Islamic University Malaysia (IIUM)
Dr. Wei Zhang, Amazon.com, Seattle, WA, USA
Mr. B. Santhosh Kumar, C S I College of Engineering, Tamil Nadu
Dr. K. Reji Kumar, N S S College, Pandalam, India

Assoc. Prof. K. Seshadri Sastry, EILM University, India
Mr. Kai Pan, UNC Charlotte, USA
Mr. Ruikar Sachin, SGGSIET, India
Prof. (Dr.) Vinodani Katiyar, Sri Ramswaroop Memorial University, India
Assoc. Prof., M. Giri, Sreenivasa Institute of Technology and Management Studies, India
Assoc. Prof. Labib Francis Gergis, Misr Academy for Engineering and Technology (MET), Egypt
Assist. Prof. Amanpreet Kaur, ITM University, India
Assist. Prof. Anand Singh Rajawat, Shri Vaishnav Institute of Technology & Science, Indore
Mrs. Hadeel Saleh Haj Aliwi, Universiti Sains Malaysia (USM), Malaysia
Dr. Abhay Bansal, Amity University, India
Dr. Mohammad A. Mezher, Fahad Bin Sultan University, KSA
Assist. Prof. Nidhi Arora, M.C.A. Institute, India
Prof. Dr. P. Suresh, Karpagam College of Engineering, Coimbatore, India
Dr. Kannan Balasubramanian, Mepco Schlenk Engineering College, India
Dr. S. Sankara Gomathi, Panimalar Engineering college, India
Prof. Anil kumar Suthar, Gujarat Technological University, L.C. Institute of Technology, India
Assist. Prof. R. Hubert Rajan, NOORUL ISLAM UNIVERSITY, India
Assist. Prof. Dr. Jyoti Mahajan, College of Engineering & Technology
Assist. Prof. Homam Reda El-Taj, College of Network Engineering, Saudi Arabia & Malaysia
Mr. Bijan Paul, Shahjalal University of Science & Technology, Bangladesh
Assoc. Prof. Dr. Ch V Phani Krishna, KL University, India
Dr. Vishal Bhatnagar, Ambedkar Institute of Advanced Communication Technologies & Research, India
Dr. Lamir LAOUAMER, Al Qassim University, Dept. Info. Systems & European University of Brittany, Dept. Computer Science, UBO, Brest, France
Prof. Ashish Babanrao Sasankar, G.H.Raisoni Institute Of Information Technology, India
Prof. Pawan Kumar Goel, Shamli Institute of Engineering and Technology, India
Mr. Ram Kumar Singh, S.V Subharti University, India
Assistant Prof. Sunish Kumar O S, Amalijothei College of Engineering, India
Dr Sanjay Bhargava, Banasthali University, India
Mr. Pankaj S. Kulkarni, AVEW's Shatabdi Institute of Technology, India
Mr. Roohollah Etemadi, Islamic Azad University, Iran
Mr. Oloruntoyin Sefiu Taiwo, Emmanuel Alayande College Of Education, Nigeria
Mr. Sumit Goyal, National Dairy Research Institute, India
Mr Jaswinder Singh Dilawari, Geeta Engineering College, India
Prof. Raghuraj Singh, Harcourt Butler Technological Institute, Kanpur
Dr. S.K. Mahendran, Anna University, Chennai, India
Dr. Amit Wason, Hindustan Institute of Technology & Management, Punjab
Dr. Ashu Gupta, Apeejay Institute of Management, India
Assist. Prof. D. Asir Antony Gnana Singh, M.I.E.T Engineering College, India
Mrs Mina Farmanbar, Eastern Mediterranean University, Famagusta, North Cyprus
Mr. Maram Balajee, GMR Institute of Technology, India
Mr. Moiz S. Ansari, Isra University, Hyderabad, Pakistan
Mr. Adebayo, Olawale Surajudeen, Federal University of Technology Minna, Nigeria
Mr. Jasvir Singh, University College Of Engg., India
Mr. Vivek Tiwari, MANIT, Bhopal, India
Assoc. Prof. R. Navaneethakrishnan, Bharathiyar College of Engineering and Technology, India
Mr. Somdip Dey, St. Xavier's College, Kolkata, India

Mr. Souleymane Balla-Arabé, Xi'an University of Electronic Science and Technology, China
Mr. Mahabub Alam, Rajshahi University of Engineering and Technology, Bangladesh
Mr. Sathyapraksh P., S.K.P Engineering College, India
Dr. N. Karthikeyan, SNS College of Engineering, Anna University, India
Dr. Binod Kumar, JSPM's, Jayawant Technical Campus, Pune, India
Assoc. Prof. Dinesh Goyal, Suresh Gyan Vihar University, India
Mr. Md. Abdul Ahad, K L University, India
Mr. Vikas Bajpai, The LNM IIT, India
Dr. Manish Kumar Anand, Salesforce (R & D Analytics), San Francisco, USA
Assist. Prof. Dheeraj Murari, Kumaon Engineering College, India
Assoc. Prof. Dr. A. Muthukumaravel, VELS University, Chennai
Mr. A. Siles Balasingh, St. Joseph University in Tanzania, Tanzania
Mr. Ravindra Daga Badgujar, R C Patel Institute of Technology, India
Dr. Preeti Khanna, SVKM's NMIMS, School of Business Management, India
Mr. Kumar Dayanand, Cambridge Institute of Technology, India
Dr. Syed Asif Ali, SMI University Karachi, Pakistan
Prof. Pallvi Pandit, Himachal Pradesh University, India
Mr. Ricardo Verschueren, University of Gloucestershire, UK
Assist. Prof. Mamta Juneja, University Institute of Engineering and Technology, Panjab University, India
Assoc. Prof. P. Surendra Varma, NRI Institute of Technology, JNTU Kakinada, India
Assist. Prof. Gaurav Shrivastava, RGPV / SVITS Indore, India
Dr. S. Sumathi, Anna University, India
Assist. Prof. Ankita M. Kapadia, Charotar University of Science and Technology, India
Mr. Deepak Kumar, Indian Institute of Technology (BHU), India
Dr. Dr. Rajan Gupta, GGSIP University, New Delhi, India
Assist. Prof. M. Anand Kumar, Karpagam University, Coimbatore, India
Mr. Mr Arshad Mansoor, Pakistan Aeronautical Complex
Mr. Kapil Kumar Gupta, Ansal Institute of Technology and Management, India
Dr. Neeraj Tomer, SINE International Institute of Technology, Jaipur, India
Assist. Prof. Trunal J. Patel, C.G.Patel Institute of Technology, Uka Tarsadia University, Bardoli, Surat
Mr. Sivakumar, Codework solutions, India
Mr. Mohammad Sadegh Mirzaei, PGNR Company, Iran
Dr. Gerard G. Dumancas, Oklahoma Medical Research Foundation, USA
Mr. Varadala Sridhar, Varadhman College Engineering College, Affiliated To JNTU, Hyderabad
Assist. Prof. Manoj Dhawan, SVITS, Indore
Assoc. Prof. Chitreshh Banerjee, Suresh Gyan Vihar University, Jaipur, India
Dr. S. Santhi, SCSVMV University, India
Mr. Davood Mohammadi Souran, Ministry of Energy of Iran, Iran
Mr. Shamim Ahmed, Bangladesh University of Business and Technology, Bangladesh
Mr. Sandeep Reddivari, Mississippi State University, USA
Assoc. Prof. Ousmane Thiare, Gaston Berger University, Senegal
Dr. Hazra Imran, Athabasca University, Canada
Dr. Setu Kumar Chaturvedi, Technocrats Institute of Technology, Bhopal, India
Mr. Mohd Dilshad Ansari, Jaypee University of Information Technology, India
Ms. Jaspreet Kaur, Distance Education LPU, India
Dr. D. Nagarajan, Salalah College of Technology, Sultanate of Oman
Dr. K.V.N.R.Sai Krishna, S.V.R.M. College, India

Mr. Himanshu Pareek, Center for Development of Advanced Computing (CDAC), India
Mr. Khaldi Amine, Badji Mokhtar University, Algeria
Mr. Mohammad Sadegh Mirzaei, Scientific Applied University, Iran
Assist. Prof. Khyati Chaudhary, Ram-eesh Institute of Engg. & Technology, India
Mr. Sanjay Agal, Pacific College of Engineering Udaipur, India
Mr. Abdul Mateen Ansari, King Khalid University, Saudi Arabia
Dr. H.S. Behera, Veer Surendra Sai University of Technology (VSSUT), India
Dr. Shrikant Tiwari, Shri Shankaracharya Group of Institutions (SSGI), India
Prof. Ganesh B. Regulwar, Shri Shankarprasad Agnihotri College of Engg, India
Prof. Pinnamaneni Bhanu Prasad, Matrix vision GmbH, Germany
Dr. Shrikant Tiwari, Shri Shankaracharya Technical Campus (SSTC), India
Dr. Siddesh G.K., : Dayananada Sagar College of Engineering, Bangalore, India
Dr. Nadir Bouchama, CERIST Research Center, Algeria
Dr. R. Sathishkumar, Sri Venkateswara College of Engineering, India
Assistant Prof (Dr.) Mohamed Moussaoui, Abdelmalek Essaadi University, Morocco
Dr. S. Malathi, Panimalar Engineering College, Chennai, India
Dr. V. Subedha, Panimalar Institute of Technology, Chennai, India
Dr. Prashant Panse, Swami Vivekanand College of Engineering, Indore, India
Dr. Hamza Aldabbas, Al-Balqa'a Applied University, Jordan
Dr. G. Rasitha Banu, Vel's University, Chennai
Dr. V. D. Ambeth Kumar, Panimalar Engineering College, Chennai
Prof. Anuranjan Misra, Bhagwant Institute of Technology, Ghaziabad, India
Ms. U. Sinthuja, PSG college of arts & science, India
Dr. Ehsan Saradar Torshizi, Urmia University, Iran
Dr. Shamneesh Sharma, APG Shimla University, Shimla (H.P.), India
Assistant Prof. A. S. Syed Navaz, Muthayammal College of Arts & Science, India
Assistant Prof. Ranjit Panigrahi, Sikkim Manipal Institute of Technology, Majitar, Sikkim
Dr. Khaled Eskaf, Arab Academy for Science ,Technology & Maritime Transportation, Egypt
Dr. Nishant Gupta, University of Jammu, India
Assistant Prof. Nagarajan Sankaran, Annamalai University, Chidambaram, Tamilnadu, India
Assistant Prof. Tribikram Pradhan, Manipal Institute of Technology, India
Dr. Nasser Lotfi, Eastern Mediterranean University, Northern Cyprus
Dr. R. Manavalan, K S Rangasamy college of Arts and Science, Tamilnadu, India
Assistant Prof. P. Krishna Sankar, K S Rangasamy college of Arts and Science, Tamilnadu, India
Dr. Rahul Malik, Cisco Systems, USA
Dr. S. C. Lingareddy, ALPHA College of Engineering, India
Assistant Prof. Mohammed Shuaib, Interat University, Lucknow, India
Dr. Sachin Yele, Sanghvi Institute of Management & Science, India
Dr. T. Thambidurai, Sun Univercell, Singapore
Prof. Anandkumar Telang, BKIT, India
Assistant Prof. R. Poorvadevi, SCSVMV University, India
Dr Uttam Mande, Gitam University, India
Dr. Poornima Girish Naik, Shahu Institute of Business Education and Research (SIBER), India
Prof. Md. Abu Kausar, Jaipur National University, Jaipur, India
Dr. Mohammed Zuber, AISECT University, India
Prof. Kalum Priyanath Udagepola, King Abdulaziz University, Saudi Arabia
Dr. K. R. Ananth, Velalar College of Engineering and Technology, India

Assistant Prof. Sanjay Sharma, Roorkee Engineering & Management Institute Shamli (U.P), India
Assistant Prof. Panem Charan Arur, Priyadarshini Institute of Technology, India
Dr. Ashwak Mahmood muhsen alabaichi, Karbala University / College of Science, Iraq
Dr. Urmila Shrawankar, G H Raison College of Engineering, Nagpur (MS), India
Dr. Krishan Kumar Paliwal, Panipat Institute of Engineering & Technology, India
Dr. Mukesh Negi, Tech Mahindra, India
Dr. Anuj Kumar Singh, Amity University Gurgaon, India
Dr. Babar Shah, Gyeongsang National University, South Korea
Assistant Prof. Jayprakash Upadhyay, SRI-TECH Jabalpur, India
Assistant Prof. Varadala Sridhar, Vidya Jyothi Institute of Technology, India
Assistant Prof. Parameshachari B D, KSIT, Bangalore, India
Assistant Prof. Ankit Garg, Amity University, Haryana, India
Assistant Prof. Rajashe Karappa, SDMCET, Karnataka, India
Assistant Prof. Varun Jasuja, GNIT, India
Assistant Prof. Sonal Honale, Abha Gaikwad Patil College of Engineering Nagpur, India
Dr. Pooja Choudhary, CT Group of Institutions, NIT Jalandhar, India
Dr. Faouzi Hidoussi, UHL Batna, Algeria
Dr. Naseer Ali Husieen, Wasit University, Iraq
Assistant Prof. Vinod Kumar Shukla, Amity University, Dubai
Dr. Ahmed Farouk Metwaly, K L University
Mr. Mohammed Noaman Murad, Cihan University, Iraq
Dr. Suxing Liu, Arkansas State University, USA
Dr. M. Gomathi, Velalar College of Engineering and Technology, India
Assistant Prof. Sumardiono, College PGRI Blitar, Indonesia
Dr. Latika Kharb, Jagan Institute of Management Studies (JIMS), Delhi, India
Associate Prof. S. Raja, Pauls College of Engineering and Technology, Tamilnadu, India
Assistant Prof. Seyed Reza Pakize, Shahid Sani High School, Iran
Dr. Thiyagu Nagaraj, University-INO, India
Assistant Prof. Noreen Sarai, Harare Institute of Technology, Zimbabwe
Assistant Prof. Gajanand Sharma, Suresh Gyan Vihar University Jaipur, Rajasthan, India
Assistant Prof. Mapari Vikas Prakash, Siddhant COE, Sudumbare, Pune, India
Dr. Devesh Katiyar, Shri Ramswaroop Memorial University, India
Dr. Shenshen Liang, University of California, Santa Cruz, US
Assistant Prof. Mohammad Abu Omar, Limkokwing University of Creative Technology- Malaysia
Mr. Snehasis Banerjee, Tata Consultancy Services, India
Assistant Prof. Kibona Lusekelo, Ruaha Catholic University (RUCU), Tanzania
Assistant Prof. Adib Kabir Chowdhury, University College Technology Sarawak, Malaysia
Dr. Ying Yang, Computer Science Department, Yale University, USA
Dr. Vinay Shukla, Institute Of Technology & Management, India
Dr. Liviu Octavian Maftciu-Scai, West University of Timisoara, Romania
Assistant Prof. Rana Khudhair Abbas Ahmed, Al-Rafidain University College, Iraq
Assistant Prof. Nitin A. Naik, S.R.T.M. University, India
Dr. Timothy Powers, University of Hertfordshire, UK
Dr. S. Prasath, Bharathiar University, Erode, India
Dr. Ritu Shrivastava, SIRTIS Bhopal, India
Prof. Rohit Shrivastava, Mittal Institute of Technology, Bhopal, India
Dr. Gianina Mihai, Dunarea de Jos" University of Galati, Romania

Assistant Prof. Ms. T. Kalai Selvi, Erode Sengunthar Engineering College, India
Assistant Prof. Ms. C. Kavitha, Erode Sengunthar Engineering College, India
Assistant Prof. K. Sinivasamoorthi, Erode Sengunthar Engineering College, India
Assistant Prof. Mallikarjun C Sarsamba Bheemna Khandre Institute Technology, Bhalki, India
Assistant Prof. Vishwanath Chikaraddi, Veermata Jijabai technological Institute (Central Technological Institute), India
Assistant Prof. Dr. Ikvinderpal Singh, Trai Shatabdi GGS Khalsa College, India
Assistant Prof. Mohammed Noaman Murad, Cihan University, Iraq
Professor Yousef Farhaoui, Moulay Ismail University, Errachidia, Morocco
Dr. Parul Verma, Amity University, India
Professor Yousef Farhaoui, Moulay Ismail University, Errachidia, Morocco
Assistant Prof. Madhavi Dhingra, Amity University, Madhya Pradesh, India
Assistant Prof. G. Selvavinayagam, SNS College of Technology, Coimbatore, India
Assistant Prof. Madhavi Dhingra, Amity University, MP, India
Professor Kartheesan Log, Anna University, Chennai
Professor Vasudeva Acharya, Shri Madhwa vadiraja Institute of Technology, India
Dr. Asif Iqbal Hajamydeen, Management & Science University, Malaysia
Assistant Prof., Mahendra Singh Meena, Amity University Haryana
Assistant Professor Manjeet Kaur, Amity University Haryana
Dr. Mohamed Abd El-Basset Matwalli, Zagazig University, Egypt
Dr. Ramani Kannan, Universiti Teknologi PETRONAS, Malaysia
Assistant Prof. S. Jagadeesan Subramaniam, Anna University, India
Assistant Prof. Dharmendra Choudhary, Tripura University, India
Assistant Prof. Deepika Vodnala, SR Engineering College, India
Dr. Kai Cong, Intel Corporation & Computer Science Department, Portland State University, USA
Dr. Kailas R Patil, Vishwakarma Institute of Information Technology (VIIT), India
Dr. Omar A. Alzubi, Faculty of IT / Al-Balqa Applied University, Jordan
Assistant Prof. Kareemullah Shaik, Nimra Institute of Science and Technology, India
Assistant Prof. Chirag Modi, NIT Goa
Dr. R. Ramkumar, Nandha Arts And Science College, India
Dr. Priyadarshini Vydhialingam, Harathiar University, India
Dr. P. S. Jagadeesh Kumar, DBIT, Bangalore, Karnataka
Dr. Vikas Thada, AMITY University, Pachgaon
Dr. T. A. Ashok Kumar, Institute of Management, Christ University, Bangalore
Dr. Shaheera Rashwan, Informatics Research Institute
Dr. S. Preetha Gunasekar, Bharathiyar University, India
Asst Professor Sameer Dev Sharma, Uttaranchal University, Dehradun
Dr. Zhihan Iv, Chinese Academy of Science, China
Dr. Ikvinderpal Singh, Trai Shatabdi GGS Khalsa College, Amritsar
Dr. Umar Ruhi, University of Ottawa, Canada
Dr. Jasmin Cosic, University of Bihac, Bosnia and Herzegovina
Dr. Homam Reda El-Taj, University of Tabuk, Kingdom of Saudi Arabia
Dr. Mostafa Ghobaei Arani, Islamic Azad University, Iran
Dr. Ayyasamy Ayyanar, Annamalai University, India
Dr. Selvakumar Manickam, Universiti Sains Malaysia, Malaysia
Dr. Murali Krishna Namana, GITAM University, India
Dr. Smriti Agrawal, Chaitanya Bharathi Institute of Technology, Hyderabad, India
Professor Vimalathithan Rathinasabapathy, Karpagam College Of Engineering, India

Dr. Sushil Chandra Dimri, Graphic Era University, India
Dr. Dinh-Sinh Mai, Le Quy Don Technical University, Vietnam
Dr. S. Rama Sree, Aditya Engg. College, India
Dr. Ehab T. Alnfwawy, Sadat Academy, Egypt
Dr. Patrick D. Cerna, Haramaya University, Ethiopia
Dr. Vishal Jain, Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), India
Associate Prof. Dr. Jiliang Zhang, North Eastern University, China
Dr. Sharefa Murad, Middle East University, Jordan
Dr. Ajeet Singh Poonia, Govt. College of Engineering & technology, Rajasthan, India
Dr. Vahid Esmaealzadeh, University of Science and Technology, Iran
Dr. Jacek M. Czerniak, Casimir the Great University in Bydgoszcz, Institute of Technology, Poland
Associate Prof. Anisur Rehman Nasir, Jamia Millia Islamia University
Assistant Prof. Imran Ahmad, COMSATS Institute of Information Technology, Pakistan
Professor Ghulam Qasim, Preston University, Islamabad, Pakistan
Dr. Parameshachari B D, GSSS Institute of Engineering and Technology for Women
Dr. Wencan Luo, University of Pittsburgh, US
Dr. Musa PEKER, Faculty of Technology, Mugla Sitki Kocman University, Turkey
Dr. Gunasekaran Shanmugam, Anna University, India
Dr. Binh P. Nguyen, National University of Singapore, Singapore
Dr. Rajkumar Jain, Indian Institute of Technology Indore, India
Dr. Imtiaz Ali Halepoto, QUEST Nawabshah, Pakistan
Dr. Shaligram Prajapat, Devi Ahilya University Indore India
Dr. Sunita Singhal, Birla Institute of Technology and Science, Pilani, India
Dr. Ijaz Ali Shoukat, King Saud University, Saudi Arabia
Dr. Anuj Gupta, IKG Punjab Technical University, India
Dr. Sonali Saini, IES-IPS Academy, India
Dr. Krishan Kumar, MotiLal Nehru National Institute of Technology, Allahabad, India

CALL FOR PAPERS

International Journal of Computer Science and Information Security

IJCSIS 2016

ISSN: 1947-5500

<http://sites.google.com/site/ijcsis/>

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc.), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embedded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at <http://sites.google.com/site/ijcsis/authors-notes> .



© IJCSIS PUBLICATION 2016

ISSN 1947 5500

<http://sites.google.com/site/ijcsis/>