

Sukkur IBA **Journal** of Computing and Mathematical Sciences

E-ISSN: 2522-3003

P-ISSN: 2520-0755

Volume: 2

No: 1

Jan - Jun

2018

Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS) is the bi-annual research journal published by **Sukkur IBA University**, Sukkur Pakistan. **SJCMS** is dedicated to serve as a key resource to provide practical information for the researchers associated with computing and mathematical sciences at global scale.

Copyright: All rights reserved. No part of this publication may be produced, translated or stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying and/or otherwise the prior permission of publication authorities.

Disclaimer: The opinions expressed in **Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)** are those of the authors and contributors, and do not necessarily reflect those of the journal management, advisory board, the editorial board, Sukkur IBA University press or the organization to which the authors are affiliated. Papers published in **SJCMS** are processed through double blind peer-review by subject specialists and language experts. Neither the **Sukkur IBA University** nor the editors of **SJCMS** can be held responsible for errors or any consequences arising from the use of information contained in this journal, instead errors should be reported directly to corresponding authors of articles.

Mission Statement

The mission of **Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)** is to provide a premier interdisciplinary platform to researchers, scientists and practitioners from the field of computing and mathematical sciences for dissemination of their findings and to contribute in the knowledge domain.

Aims & Objectives

Sukkur IBA Journal of Computing and Mathematical Sciences aims to publish cutting edge research in the field of computing and mathematical sciences.

The objectives of **SJCMS** are:

1. to provide a platform for researchers for dissemination of new knowledge.
2. to connect researchers at global scale.
3. to fill the gap between academician and industrial research community.

Research Themes

The research focused on but not limited to following core thematic areas:

Computing:

- Software Engineering
- Formal Methods
- Human Computer Interaction
- Information Privacy and Security
- Computer Networks
- High Speed Networks
- Data Communication
- Mobile Computing
- Wireless Multimedia Systems
- Social Networks
- Data Science
- Big data Analysis
- Contextual Social Network Analysis and Mining
- Crowdsourcing Management
- Ubiquitous Computing

- Distributed Computing
- Cloud Computing
- Intelligent devices
- Security, Privacy and Trust in Computing and Communication
- Wearable Computing Technologies
- Soft Computing
- Genetic Algorithms
- Robotics
- Evolutionary Computing
- Machine Learning

Mathematics:

- Applied Mathematical Analysis
- Mathematical Finance
- Applied Algebra
- Stochastic Processes

Patron's Message

Sukkur IBA University has been imparting education with its core values merit, quality, and excellence since its inception. Sukkur IBA University has achieved numerous milestones in a very short span of time that hardly any other institution has achieved in the history of Pakistan. The university is continuously being ranked as one of the best university in Pakistan by Higher Education Commission (HEC). The distinct service of Sukkur IBA University is to serve the rural areas of Sindh and also underprivileged areas of other provinces of Pakistan. Sukkur IBA University is committed to serve targeted youth of Pakistan who is suffering from poverty and deprived of equal opportunity to seek quality education. Sukkur IBA University is successfully undertaking its mission and objectives that lead Pakistan towards socio-economic prosperity.

In continuation of endeavors to touch new horizons in the field of computing and mathematical sciences, Sukkur IBA University publishes an international referred journal. Sukkur IBA University believes that research is an integral part of modern learnings and development. Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS) is the modest effort to contribute and promote the research environment within the university and Pakistan as a whole. SJCMS is a peer-reviewed and multidisciplinary research journal to publish findings and results of the latest and innovative research in the fields, but not limited to Computing and Mathematical Sciences. Following the tradition of Sukkur IBA University, SJCMS is also aimed at achieving international recognition and high impact research publication in the near future.

Prof. Nisar Ahmed Siddiqui

(Sitara-e-Imtiaz)

Vice Chancellor, Sukkur IBA University

Patron SJCMS

Publisher: Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)

Office of Research Innovation & Commercialization – ORIC

Sukkur IBA University – Airport Road Sukkur-65200, Sindh Pakistan

Tel: (092 71) 5644233 Fax: (092 71) 5804425 Email: sjcms@iba-suk.edu.pk URL: sjcms.iba-suk.edu.pk

Editorial

Dear Readers,

It is pleasure to present to you the first issue of volume 2 of Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS).

In today's globalized, interconnected world, information and computational sciences provide variety of tools that have great potential to contribute for creating a prosperous and inclusive society. The modern tools and techniques to solve the complex problems of the world we living in, are the outcome of advances and innovations in various fields of computing where the mathematical science remains at the core of all developments. The SJCMS aims to publish cutting-edge research in the field of computing and mathematical sciences for dissemination to the largest stakeholders. SJCMS has achieved milestones in very short span of time and is indexed in renowned databases such as DOAJ, Google Scholar, DRJI, BASE, ROAD, CrossRef and many others.

This issue contains the double-blind peer-reviewed articles that address the key research problems in the specified domain The SJCMS adopts all standards that are a prerequisite for publishing high-quality research work. The Editorial Board and the Reviewers Board of the Journal is comprised of renowned researchers from technologically advanced countries. The Journal has adopted the Open Access Policy without charging any publication fees that will certainly increase the readership by providing free access to a wider audience.

On behalf of the SJCMS, I welcome the submissions for upcoming issue (Volume-2, Issue-2, July-December 2018) and looking forward to receiving your valuable feedback.

Sincerely,

Ahmad Waqas, PhD
Chief Editor

Publisher: **Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)**

Office of Research Innovation & Commercialization – ORIC

Sukkur IBA University – Airport Road Sukkur-65200, Sindh Pakistan

Tel: (092 71) 5644233 Fax: (092 71) 5804425 Email: sjcms@iba-suk.edu.pk URL: sjcms.iba-suk.edu.pk

*Patron***Prof. Nisar Ahmed Siddiqui***Chief Editor***Dr. Ahmad Waqas***Associate Editors***Dr. M. Abdul Rehman, Dr. Javed Hussain Brohi***Managing Editors***Prof. Dr. Pervez Memon, Dr. Sher Muhammad Daudpota***Editorial Board***Prof. Dr. Abdul Majeed Siddiqui**
Pennsylvania State University, USA**Prof. Dr. Gul Agha**
University of Illinois, USA**Prof. Dr. Muhammad Ridza Wahiddin**
International Islamic University, Malaysia**Prof. Dr. Tahar Kechadi**
University College Dublin, Ireland**Prof. Dr. Paolo Bottoni**
Sapienza - University of Rome, Italy**Prof. Dr. Md. Anwar Hossain**
University of Dhaka, Bangladesh**Dr. Umer Altaf**
KAUST, Kingdom of Saudi Arabia**Prof. Dr. Farid Nait Abdesalam**
Paris Descartes University Paris, France**Prof. Dr. Asadullah Shah**
International Islamic University, Malaysia**Prof. Dr. Adnan Nadeem**
Islamia University Madina, KSA**Dr. Jafreezal Jaafar**
Universiti Teknologi PETRONAS**Dr. Zulkefli Muhammad Yusof**
International Islamic University, Malaysia**Dr. Hafiz Abid Mahmood**
AMA International University, Bahrain**Prof. Dr. S.M Aqil Burney**
IoBM, Karachi, Pakistan**Prof. Dr. Zubair Shaikh**
Muhammad Ali Jinnah University, Pakistan**Prof. Dr. Mohammad Shabir**
Quaid-i-Azam University Islamabad, Pakistan**Dr. Ferhana Ahmad**
LUMS, Lahore, Pakistan**Dr. Asghar Qadir**
Quaid-e-Azam University, Islamabad**Dr. Nadeem Mahmood**
University of Karachi, Pakistan**Engr. Zahid Hussain Khand**
Sukkur IBA University, Pakistan**Dr. Qamar Uddin Khand**
Sukkur IBA University, Pakistan**Dr. Syed Hyder Ali Muttaqi Shah**
Sukkur IBA University, Pakistan**Dr. Muhammad Ajmal Sawand**
Sukkur IBA University, Pakistan**Dr. Niaz Hussain Ghumro**
Sukkur IBA University, Pakistan**Dr. Zarqa Bano**
Sukkur IBA University, Pakistan**Dr. Javed Ahmed Shahani**
Sukkur IBA University, Pakistan*Language Editor - Prof. Ghulam Hussain Manganhar*

Publisher: Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)**Office of Research Innovation & Commercialization – ORIC****Sukkur IBA University – Airport Road Sukkur-65200, Sindh Pakistan**Tel: (092 71) 5644233 Fax: (092 71) 5804425 Email: sjcms@iba-suk.edu.pk URL: sjcms.iba-suk.edu.pk

Contents

Title	Pages
Genetic Algorithm-based Optimized Fuzzy Adaptive Path Selection in Wireless Sensor Networks <i>Muhammad Akram, Muhammad Ashraf, Tae Ho Cho</i>	(1-12)
Crime Mapping in GIS by Using Hotspot <i>Syeda Ambreen Zahra</i>	(13-19)
Evaluation of Android Malware Detectors <i>Hassan Rafiq, Muhammad Aleem, Muhammad Arshad Islam</i>	(20-29)
Challenges of Computer Science and IT in Teaching-Learning in Saudi Arabia <i>Hafiz Abid Mahmood Malik, Faiza Abid, R. Kalaicelvi, Zeeshan Bhatti</i>	(30-37)
Holy Qur'an Speech Recognition System Distinguishing the Type of prolongation <i>Bilal Yousfi, Akram M. Zeki, Aminah Haji</i>	(38-45)
Analysis on Energy Efficient Protocols-Wireless Sensor Networks <i>Iqra Tariq, Talal Bin Maqsood, Dr. Babur Hayat Malik, Mareena Asghar, Quratulain Gulzar</i>	(46-54)
Analysis on Security Methods of Wireless Sensor Network (WSN) <i>Murtaza Ahmed Siddiqi, Abdul Aziz Mugheri, Mohammad Khoso</i>	(55-63)
Swarm Based Coverage Using Multiple Informed Leaders <i>Ahmad Din, Ashfaq Ahmed, Kashif Zia, Abbas Khalid, Owais Khan</i>	(64-72)
Comparative Study of Testing Tools Blazemeter and Apache Jmeter <i>Pirah Memon, Tahseen Hafiz, Sania Bhatti, Saman Shahid Qureshi</i>	(73-80)
Internet of Things (IoTs) for Disaster Management <i>Syeda Ambreen Zahra, Iqra Shafique, Tuba Farid</i>	(81-89)

Publisher: Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)

Office of Research Innovation & Commercialization – ORIC

Sukkur IBA University – Airport Road Sukkur-65200, Sindh Pakistan

Tel: (092 71) 5644233 Fax: (092 71) 5804425 Email: sjcms@iba-suk.edu.pk URL: sjcms.iba-suk.edu.pk

Contents

- Genetic Algorithm-based Optimized Fuzzy Adaptive Path Selection in Wireless Sensor Networks (1-12)
Muhammad Akram, Muhammad Ashraf, Tae Ho Cho
- Crime Mapping in GIS by Using Hotspot (13-19)
Syeda Ambreen Zahra
- Evaluation of Android Malware Detectors (20-28)
Hassan Rafiq, Muhammad Aleem, Muhammad Arshad Islam
- Challenges of Computer Science and IT in Teaching-Learning in Saudi Arabia (29-35)
Hafiz Abid Mahmood Malik, Faiza Abid, R. Kalaicelvi, Zeeshan Bhatti
- Holy Qur'an Speech Recognition System Distinguishing the Type of prolongation (36-43)
Bilal Yousfi, Akram M. Zeki, Aminah Haji
- Analysis on Energy Efficient Protocols-Wireless Sensor Networks (44-51)
Iqra Tariq, Talal Bin Maqsood, Dr. Babur Hayat Malik, Mareena Asghar, Quratulain Gulzar
- Analysis on Security Methods of Wireless Sensor Network (WSN) (52-60)
Murtaza Ahmed Siddiqi, Abdul Aziz Mugheri, Mohammad Khoso
- Swarm Based Coverage Using Multiple Informed Leaders (61-69)
Ahmad Din, Ashfaq Ahmed, Kashif Zia, Abbas Khalid, Owais Khan



Vol. 2, No. 1 | Jan – June 2018



Comparative Study of Testing Tools Blazemeter and Apache Jmeter (70-76)
Pirah Memon, Tahseen Hafiz, Sania Bhatti, Saman Shahid Qureshi

Internet of Things (IoTs) for Disaster Management (77-85)
Syeda Ambreen Zahra, Iqra Shafique, Tuba Farid

A Genetic Algorithm-based Optimized Fuzzy Adaptive Path Selection in Wireless Sensor Networks

Muhammad Akram¹, Muhammad Ashraf¹, Tae Ho Cho²

Abstract:

In Wireless sensor networks, energy efficiency can be achieved by adaptive choice of the data forwarding path to balance the energy dissipation in the network. This adaptive path selection is done through a fuzzy rule-based method given the input parameters. Due to uncertainty in reasoning and inferencing process and imprecision in the data, the fuzzy-based system becomes an ideal choice for the selection of the paths. In fuzzy systems, the membership functions need to be optimized to make the best use of the fuzzy inferencing and improve the performance of the fuzzy system. Genetic algorithm-based fuzzy membership function optimization technique selects the optimal solution in a feasible time and saves from the hassle of manual intervention. Manual optimization efforts are unfeasible for common applications and take unlimited time and human expertise to optimize functions in an exhaustive search field. This technique assesses the fitness of the membership functions through simulation outcomes and optimizes them through genetic algorithm based evaluation process. The proposed scheme consists of three modules; The first module simulates the membership function in the given network model, the second module analyzes the performance efficiency of the membership functions through simulation, and the last module constructs the subsequent membership-function populations using GA techniques. The proposed method automatically optimizes the membership functions in the fuzzy system with little human intervention, requires minimal human expertise and saves ample time in the optimization process.

Keywords: *Fuzzy optimization; Route selection; Filtering; Genetic algorithm; Fuzzy.*

1. Introduction

Wireless sensor networks are economically a feasible tool for monitoring physical world, usually comprising of a many number of sensor nodes, and they are often densely installed in hostile locations [1]-[4]. Sensor networks are expected to function as a flexible, universal and effortlessly installable solution for cyber-physical systems and applications [5]. Sensor nodes are usually placed unattended in exposed surroundings, therefore, they are extremely susceptible to being physically captured and compromised [6]. Nodes in the network can be compromised and exploited by the attacker to insert false sensing

information into the network, which both causes false alarms at the base station (BS) and waste the scarce energy resources of the data forwarding nodes in the network [7]. Similarly, compromised nodes are misused by the attackers to insert counterfeit message authentication codes (MACs), alternatively referred to as votes, to cause the dropping of the legitimate reports at the intermediate nodes during the en-route verification process [4], [7]. Injection of false votes restricts the true information from reaching the BS. As show in Fig. 1, an attacker can exploit a compromised node to either inject fabricated data with false votes or attach false votes to legitimate data to

¹ College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea.

² College of Software, Sungkyunkwan University, Suwon 16419, Republic of Korea

Corresponding Email: akram.khan@skku.edu

cause the aforementioned threats to the network. Researchers have suggested various security schemes to filter fabricated data en-route, and deliver the legitimate information to the BS given the number of fake votes attached to a legitimate report is less than a certain value [3], [4], [7]-[9].

PVFS and FASIN filter fabricated reports en-route before they consume significant energy resource of the en-route nodes [4], [7].

PVFS and FASIN also allow true reports with false MACs, which are less than certain number, to be delivered to the BS. In PVFS, every cluster member node share their own authentication keys with the verification nodes probabilistically whereas FASIN employ a fuzzy inferencing system installed at each CH which helps to adaptive select the verification on the data forwarding paths [3], [7]. In minimum cost forwarding-based (MCF) schemes [[10], routing paths are chosen by only considering the distance and the energy efficiency of the routing paths. Constant uses of shortest or minimum cost paths leads to uneven work load across the network and causes network partitioning [11]. In [11], we recommended fuzzy based adaptive selection of data routing path to facilitate energy balancing across the network and avoid network partitioning. Fuzzy adaptive choice of data routing paths feasibly balance the work burden between different available routes. Load balancing spreads the utilization of the energy resources in the network, potentially leading to an extended network lifetime [12]. The verification nodes on all the available paths are equal in number which implies that the filtering capability of all the routing paths is essentially same. The adaptive filtering route selection achieves extra energy saving and results in network lifetime extension. In fuzzy adaptive path selection (shortly referred to as FAPS hereafter) [11], a fuzzy rule-based inferencing system chooses among the available data forwarding paths, considering the input parameters such as the verification nodes' average distance, the distance of the

routing path, and the average residual energy resources on the data routing path. To exploit FAPS in a real-world WSN, the fuzzy membership functions in the fuzzy inferencing configuration are required to be optimally tuned for the actual network configuration. However, such an optimization is usually beyond human skills [13]. Even if it is likely to manually optimize the system, the time-cost of manual optimization would be unpractical for real-world WSN configurations.

With prior knowledge it is easy to understand and construct the IF-THEN structure of fuzzy rules. However, many parameters are specified by experts and the identification of these parameters can be treated as optimization problem. Several optimization techniques have been proposed. However, all these techniques do not always guarantee the finding of an optimal solution in multi-parametric space [13]. A genetic algorithm based membership function optimization technique has proved to be more reliable in discovering the ideal solution in multi-parametric space [13]. A genetic algorithm (GA) is a concurrent, global search method that imitates natural genetic processes [14]. GA becomes an ideal choice for membership function optimization due to its ability to simultaneously explore multiple points in the search space and it converge to a near optimal solution. The search space does not need to be necessarily differentiable and continuous for the GA to work, and it has the capability to iterate several times on each piece of received data [13].

2. Fuzzy Adaptive Path Selection (FAPS)

In [11] we presented a fuzzy rule based adaptive path selection scheme (FAPS) which also makes use of fuzzy adaptive selection of verification nodes (FASIN) [4]. In FAPS, each intermediate path between the source cluster and the base station (BS) have equal number of verification node which are initially selected probabilistically. FASIN [4] use fuzzy rule-based inferencing to select verification nodes

on each path and gradually improves their average distance to the source cluster. FAPS selects the suitable routing path considering the following input parameters for each path.

1. The number of verification nodes with in L hops on the path ($L =$ cluster size)
2. The path distance
3. The average residual energy of the nodes on the path.

In FAPS, the membership functions of the input and output parameters are carefully selected to obtain desirable results based on the simulation outputs. These membership functions can be modified and adjusted based on the professional understanding of the network parameters such as network size and

density of nodes. However, in real-world WSNs, manual optimization of the input membership functions requires a lot of efforts and expert knowledge and is mostly impractical as it may not obtain an optimal solution within reasonable time. Such an optimization task is usually beyond human expertise and capabilities because the space has to be exhaustively searched for the optimal solution.

We propose to optimize membership functions for FAPS in WSNs using genetic algorithms. The primary goal of GA-based fuzzy adaptive path selection (GAFPS) is to provide fuzzy membership functions that are optimized and function well in accordance with the network configurations requiring little human expertise and involvement.

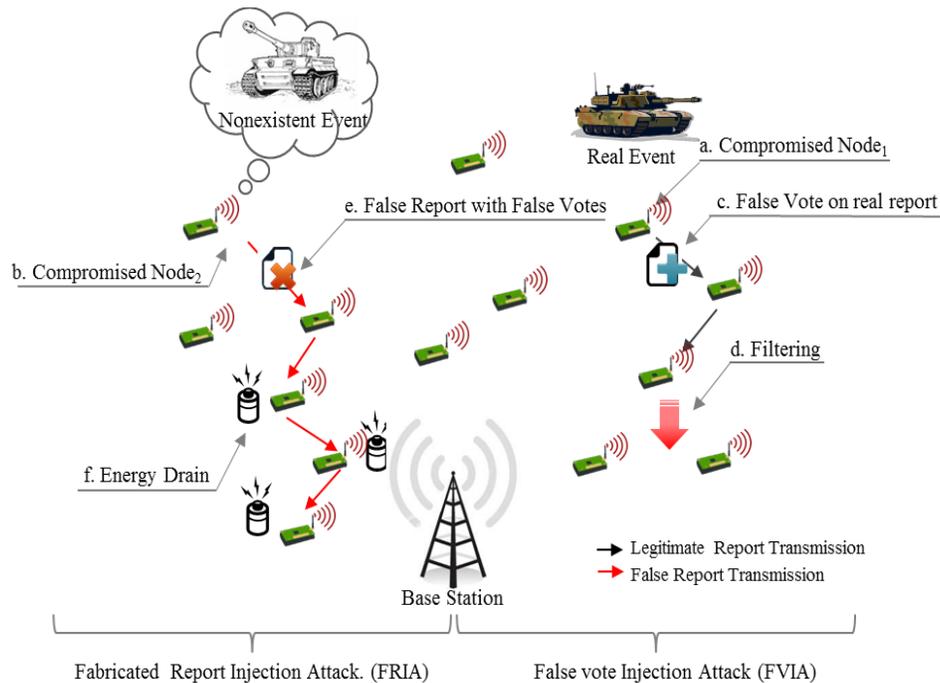


Fig. 1. False report and false vote injection attacks in wireless sensor network.

3. GA-based Fuzzy adaptive path selection (GAFPS)

This section describes the Genetic algorithm-based optimization of the fuzzy input parameters for the fuzzy adaptive selection of data routing paths.

3.1. Genetic Representation of Membership Functions

In GA-based membership function optimization scheme, a single chromosome is used to represent the set of fuzzy input membership functions during a single trial.

The input parameters of the fuzzy rule based system in [11] are:

1. REL
2. NKIH
3. DIST

The output parameter of the fuzzy rule based system is:

1. Fitness

Each input parameter has 3 triangular membership functions which are labeled as follows:

1. REL: L (Low) , M (Medium), H (High)
2. NKIH: Le (Less), Mo (Moderate), Mr (More)
3. DIST: N (Near), F (Fair), Fa (Far)

Therefore the fuzzy rule-based system makes use of 27 ($3 \times 3 \times 3$) fuzzy rules. Similarly, the output of the fuzzy rule-based system produces the inferred fitness value of the path for which the fuzzy inferencing is being carried out. The output parameter consists of 4 triangular membership function denoted by:

1. Fitness: P (Poor), Mo (Moderate), B (Better), Be (Best)

The design of the membership functions should be such that they satisfy two conditions given below:

- a. Every membership function intersects only with the nearest adjacent membership functions;
- b. Membership values should be equal to or nearly 1 in related fuzzy sets.

Membership function can be conveniently represented through chromosomes. Chromosomes are binary coded comprising of bits: 1s and 0s. Each bit in a chromosome represents a gene value which is either 0 or 1. A single chromosome represents a single trial set of input fuzzy membership functions as depicted in Fig. 2. One parameter suffices to represent 3 membership functions of triangular shape for a single input. Therefore, 3 parameters for the input membership functions are coded into a chromosome for each input variable for its representation in a chromosome. Each parameter of REL, DIST are coded by 6 binary digits whereas the single parameter of variable NKIH is coded by 3 binary digits. Therefore, a single chromosome size is $(6+6+3) = 15$ bits.

The optimization process is carried out in design-time prior to network deployment. Exhaustive search strategy may not find an optimum solution within bounded time; it takes $2^{15}(=32768)$ trial sets and each set is simulated for its performance evaluation. We don't modify the fuzzy rules and they are fixed. Modifying the fuzzy rules also widens the search space which increases the time cost exponentially, and we cannot expect a near optimal solution [15]. In GAFPS, 300 generations, each comprising a population of 30 chromosomes, are created. Therefore, only 9,000 trial sets are tested in simulations. Fig. 2 shows the representation of the input fuzzy membership function through a single chromosome.

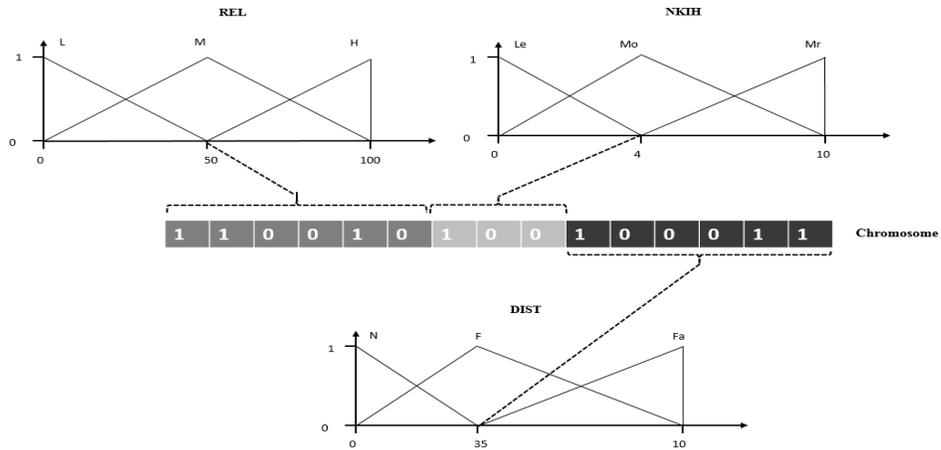


Fig. 2. Genetic representation of membership functions.

3.2. GA Module

The GA module stores a collection of chromosomes that are represented by a sequence of binary strings. Initially, the bit values in a single chromosome are set randomly with a probability value of 0.5 for setting each bit in the chromosome. Each chromosome is tested through simulations for their fitness values. Subsequent generations of chromosome are chosen according to evolutionary techniques employed in genetic algorithms such as selection, crossover, mutation etc. Fig. 3 show the evolution process and explains the construction of the subsequent generations of chromosomes.

A random pair of chromosomes is chosen for reproduction in the next generation according to their probabilities. Genetic operator such as selection, crossover and mutation help to boost chromosome up towards the optimum. Selection realizes the survival of the fittest in the evolution process. Fig. 3 shows the evolution of a 4-parameter (for example) input fuzzy membership function. Fig. 3(a) illustrates the selection of a chromosome in the i^{th} generation to be reproduced in the $(i+1)^{\text{th}}$ generation. GAFPS uses the fitness proportional method for the selection of a pair of chromosomes to reproduce their next generation offspring. In fitness proportional method, the selection

probability of a chromosome for reproduction is directly proportional to its fitness value. The chromosome, whose fitness value is greater, is highly likely to be selected, therefore the average fitness value of the entire population improves over generations. A chromosome with the higher fitness value can be selected more than once whereas the chromosome with the least fitness value may not be chosen at all. In Fig. 3(a), chromosome c4 has the highest fitness value and has been selected twice whereas chromosome c2, with the lowest fitness value, has not been selected at all. Genetic operator, cross over, imitates the genetic inheritance in the offspring and recombines segments of the individuals corresponding to parents. It ensures the exploration of search space.

Fig. 3(b) depicts the crossover operation. Crossover produces a swap or a shift of the fractional parameters to produce new membership functions as depicted in Fig. 3(b) and Fig. 3(c), respectively. Fig. 3(b) shows the interchange of a single parameter between $c1'$ and $c2'$, $c2'$ and $c3'$, and $c3'$ and $c4'$ whereas Fig. 3(c) shows the shift in the membership function achieved through partial altering of the membership function. The altering of the parameter can also be done through mutation which imitates the genetic mutation (self-variation) of genes. The mutation operator is

another means of exploring search space and only a small number of genes (bits) must be altered in a chromosome in accordance with a designed probability. Fig. 3 (d) shows the mutation operation wherein few bits of the chromosome c^j are partially altered to produce a new membership function.

Optimization process can be done more than once to avoid convergence to local optima/maxima. Therefore, we have carried out the optimization process twice changing the crossover and mutation probability values. In the first round of optimization, probability values of 0.9 for crossover and 0.01 for mutation are used. In the second round of optimization, the crossover and mutation probability values are changed to 0.6 and 0.1 respectively. GAFPS selects the chromosome that has a greater fitness value.

3.3. Simulation Module

During each generation, every chromosome is evaluated for its fitness value through a simulation run in the simulation module. Simulation module simulates the target WSN for each single chromosome. The Simulation module comprises of two models i.e. a network model and an attack model. Fig. 4 depicts the simulation process in the simulation module. When a chromosome arrives at the simulation module, it breaks it into 3 segments and reconstructs the fuzzy sets. The fuzzy membership functions are supplied to the rule based system of the FAPS as in [11] in the simulation module. The fitness evaluation module uses results of each simulation run to calculate the individual chromosome's fitness value. During each simulation run, we measure the residual energy of the network, the average number of hop traversed by a report before a first false MAC is detected, and the average number of hops traversed by false reports before they are filtered.

3.4. Fitness Evaluation Module

Fitness evaluation module calculates the fitness of every individual chromosome using the results produced by the simulation unit. A chromosome c_j 's fitness value $F_{c_{ij}}$ is calculated by the following equation:

$$F_{c_{ij}} = (\alpha \times (1/n \sum_{i=1}^n REL_i)) + (\beta \times \frac{1}{H}) \quad (1)$$

In the above equation, c_{ij} is the i^{th} chromosome in the j^{th} generation, REL = the residual energy of the network, n = size of the network, H = the average number of hops traversed by false reports, and α and β are weight factors.

The fitness value calculated by the fitness evaluation module is provided to the GA module. The value of the weight factors are determined with regards to the objective of the target WSN. Since, the energy of the network is the most important factor, therefore value of α is usually much greater than the values of β [16]. Since the detection power of all the paths is same in terms of the number of verification nodes, the drop rate of false reports and delivery success rate of legitimate reports are irrelevant and not considered in calculating the fitness value of a chromosome. We are only interested in improving the energy saving of the sensor network and early filtering of the false reports.

3.5. Optimized Fuzzy Adaptive Path Selection

The GA module contains a population of 30 chromosomes in each generation. The simulation runs and regenerates a new population using genetic operations until the terminal condition reaches i.e. 300 generations. Total 400 reports are generated during each simulation run including legitimate reports with true MACs, legitimate reports with false MACs, and false reports. Fig. 5 illustrates the optimized input fuzzy membership functions.

A randomly deployed static network with 4000 homogenous nodes is simulated in a custom simulator developed in Microsoft C++ 2012. The network is contained in a $1200 \times 1000 \text{ m}^2$ 2-D terrain. The cluster size is $L=10$, and every node in the cluster possesses a single key for authenticating the report. If the threshold value T_f reaches 3, the report will immediately be dropped, whereas a report is accepted as a legitimate report if the value of T_f reaches 4. Each reports is authenticated by 5 nodes randomly picked up in the cluster

including CH. 16.25/12.5 μJ of energy per byte are consumed to transmit/receive, and 15 μJ are consumed to generate a vote. The size of the report and MAC is 36 and 4 bytes, respectively. MAC generation consumes 15 μJ of energy. The verification of a report at a

particular node consumes 75 μJ of energy. The BS is located at the upper left edge of the network. The network is exposed only to FVIA and FRIA attacks. Cluster based organization restricts compromised nodes in different clusters from colluding with one another.

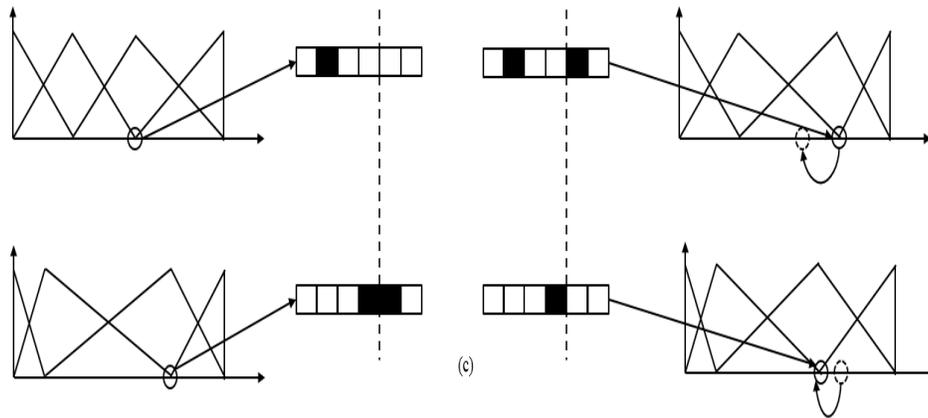
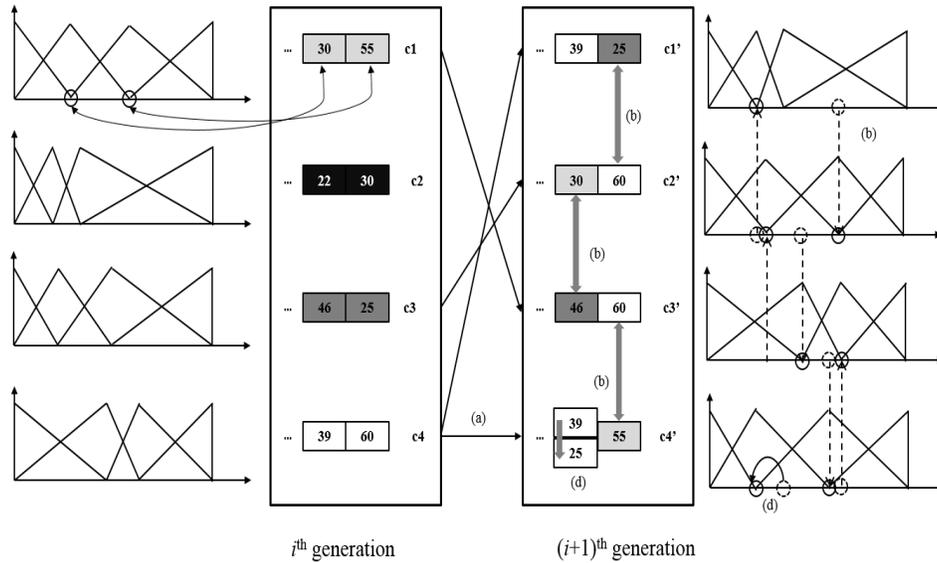


Fig. 3. Evaluation of next generation of chromosome population.

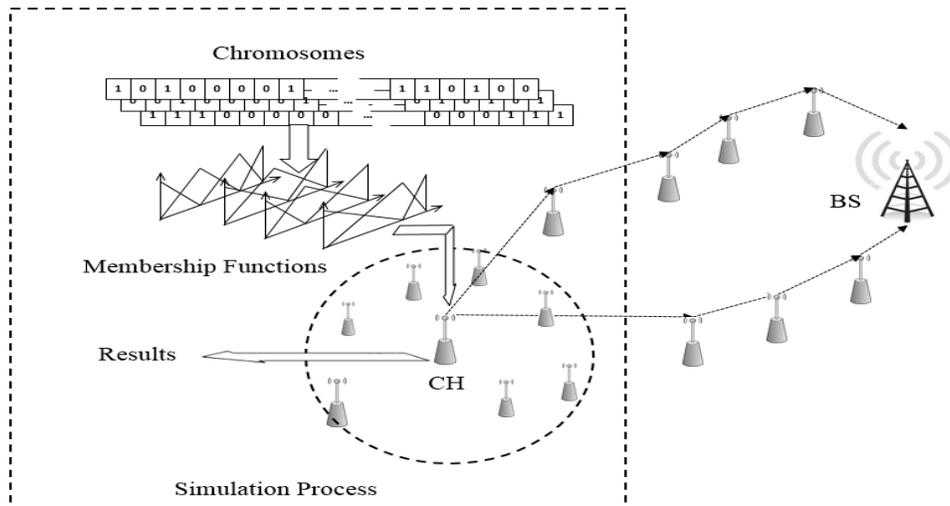


Fig. 4. Simulation Process.

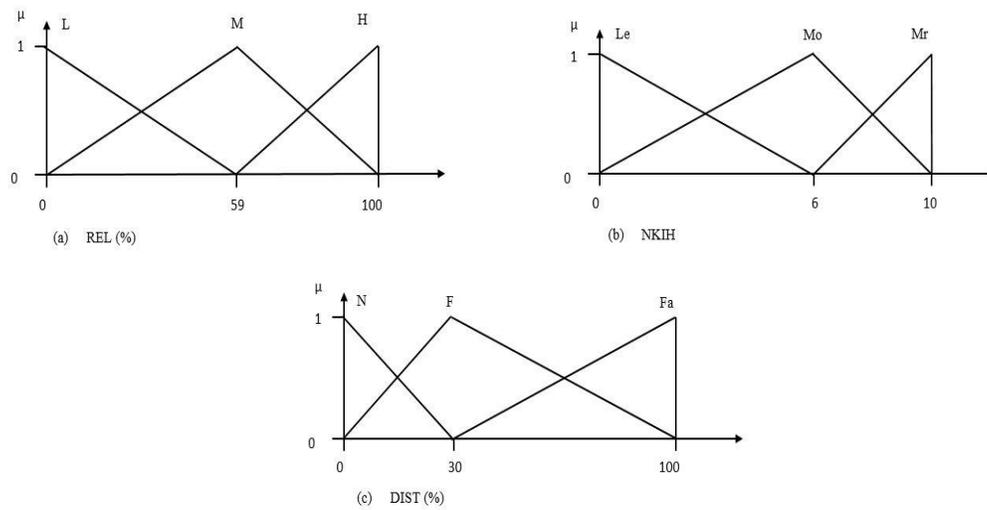


Fig. 5. Optimized input fuzzy membership functions.

4.Simulation Results

Fig. 6 shows the average energy invested on a false report being forwarded to the BS before getting filtered during different generations. As illustrated in Fig. 6, the average energy consumption per filtered report in 1st and 40th generations is greater than in the Final generation. The reason for higher energy consumption in 1st and 40th generation is that the membership functions are chosen such that it may choose a path with farther verification nodes than a path which has more nodes closer to the source cluster.

An arbitrary and haphazard design of membership functions chooses a less suitable path, and hence an increase in energy dissipation, and irregular pattern in the energy curve. This analysis reveals that the careful selection of membership functions significantly impacts the performance of the network. Since the membership function are optimized in the final generation therefore, it exhibits a steady behavior and the energy consumption is lesser than in non-optimized generations. The energy consumption of the final generation compares well with that of the

manually optimized solution and in fact save further energy.

Fig. 7 shows the improvement in the average fitness value of the chromosomes in the population during each generation. The values of the chromosomes are generated randomly during the 1st generation and GA technique such as crossover, mutation and selection are applied in each generation to improve the average fitness of the population. As shown, the average fitness value of the population improves and eventually stabilizes.

Fig. 8 shows that the optimization of fuzzy membership functions also helps to improve the performance of the proposed scheme with regard to the average number of hops traversed by a report before the first false MAC is detected. Optimized membership functions ensure the selection of the appropriate and most feasible route for data forwarding. Therefore, the energy saving capability of the network also improves and false reports are filtered relatively earlier.

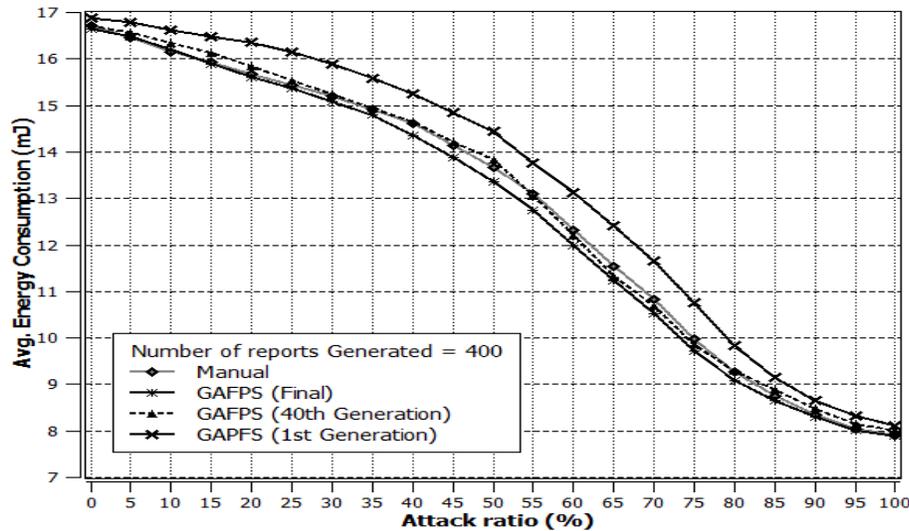


Fig. 6. Average Energy consumption per report before getting filtered.

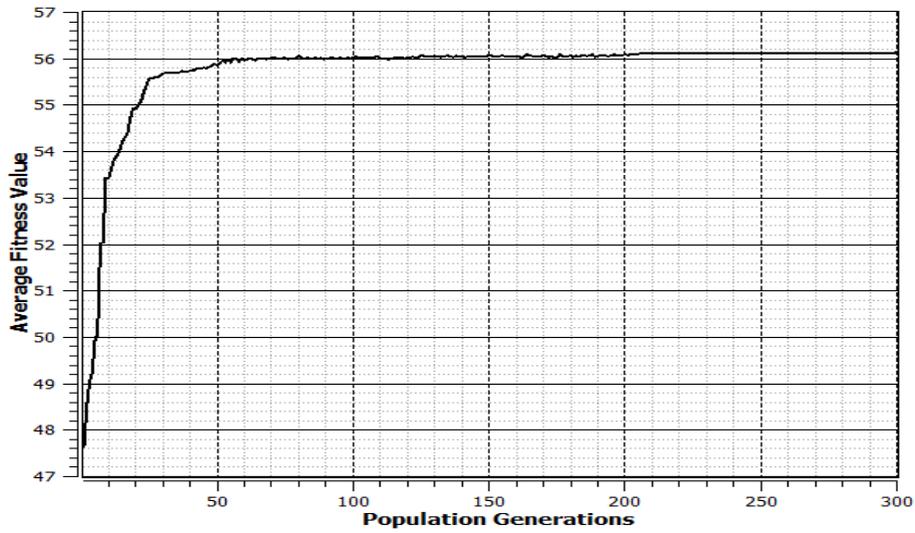


Fig. 7. Average fitness value.

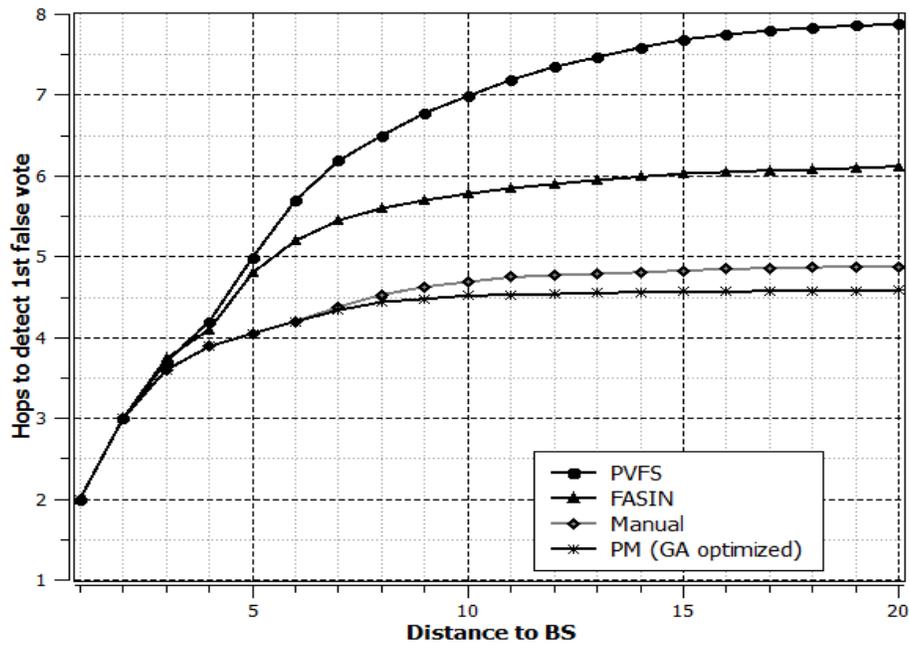


Fig. 8. Hops travelled by a report until first false MAC is detected.

5. Time Cost of Fuzzy Optimization

An exhaustive search method for the optimization of fuzzy membership functions may literally take several years usually in order of thousand years [17]. GA based membership-function optimization technique saves time cost and retrieves the optimal solution in much lesser time. The GAFPS finds an optimal solution for 9,000 trial sets (30 chromosomes \times 300 generations) in less than a day in the worst case i.e., if all the 9,000 trial sets are different from one another. Furthermore, during the evaluation of the generations, several identical chromosomes exist in a particular generation of chromosomes. Therefore further energy can be saved by evaluating only one of the identical chromosomes. Similarly, in the subsequent generations, several offspring chromosomes are identical to their parent chromosomes in the previous generations since the crossover and mutation probabilities are usually less than 1. The probability of having identical chromosomes in a single generation increases with an increasing number of identical chromosomes in a single population as the optimization process progresses. Therefore GAFPS finds an optimal solution within a feasible time for the network.

6. Conclusion

Genetic algorithms are an excellent tool for optimization of fuzzy membership function and require little human expertise. In this paper, we use GA based optimization technique to fine-tune the fuzzy membership functions in FAPS. GA based membership-function optimization technique saves time cost and retrieves the optimal solution in much lesser time. A chromosome represents the parameters that define the membership functions. Each chromosome in every generation is evaluated for its performance in a simulation unit. Genetic operators such as crossover, mutation are applied with their designed probabilities during the evaluation process and selection operator chooses the fittest chromosomes in the current population to be reproduced in the next population. Network energy, average hops traversed by

false reports are the two factors that are used in evaluating a chromosome for its fitness value. A significant amount of energy is saved by evaluating only one of the identical chromosomes. GA based optimization processes that make use of natural evolution methods, present a promising tool that supports the optimization of the parameters of fuzzy rule-based systems in WSNs.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. NRF-2015R1D1A1A01059484).

REFERENCES

- [1] L. Buttyán, L. Dóra, and I. Vajda, "Statistical wormhole detection in sensor networks," in *European Workshop on Security in Ad-Hoc and Sensor Networks*, pp. 128-141, 2005.
- [2] X. Han et al, "Fault-tolerant relay node placement in heterogeneous wireless sensor networks," *IEEE Transaction on Mobile Computing*, vol. 9, no. 5, pp. 643-656, 2010.
- [3] F. Ye et al, "Statistical en-route filtering of injected false data in sensor networks," *IEEE journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 839-850, 2005.
- [4] M. Akram and T. H. Cho, "Energy efficient fuzzy adaptive selection of verification nodes in wireless sensor networks," *Ad Hoc Networks*, vol. 47, pp. 16-25, 2016.
- [5] A. Wood, V. Srinivasan, and J. Stankovic, "Autonomous defenses for security attacks in pervasive CPS infrastructure," in *Proc. DHS: S&T Workshop on Future Directions in Cyber-Physical Systems Security*, 2009.
- [6] H. Y. Lee and T. H. Cho, "Fuzzy adaptive selection of filtering schemes for energy saving in sensor networks," *IEICE Trans. Commun.*, vol. 90, no. 12, pp. 3346-3353, 2007.
- [7] F. Li, A. Srinivasan, and J. Wu, "PVFS: a probabilistic voting-based filtering scheme in wireless sensor networks," *International Journal of Security and Networks*, vol. 3, no. 3, pp. 173-182, 2008.

- [8] Z. Yu and Y. Guan, "A dynamic en-route scheme for filtering false data injection in wireless sensor networks," in *SenSys*, 2005, pp. 294-295.
- [9] S. Zhu et al, "An interleaved hop-by-hop authentication scheme for filtering of injected false data in sensor networks," in *Proceedings of IEEE symposium on Security and Privacy*, 2004, pp. 259-271.
- [10] F. Ye et al, "A scalable solution to minimum cost forwarding in large sensor networks," in *Computer Communications and Networks, 2001. Proceedings. Tenth International Conference On*, 2001, pp. 304-309.
- [11] M. Akram and T. H. Cho, "Energy Efficient Fuzzy Adaptive Verification Node Selection-Based Path Determination in Wireless Sensor Networks," *Symmetry*, vol. 9, no. 10, pp. 220, 2017.
- [12] D. Ganesan et al, "Highly-resilient, energy-efficient multipath routing in wireless sensor networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 4, pp. 11-25, 2001.
- [13] J. Kim, Y. Moon, and B. P. Zeigler, "Designing fuzzy net controllers using GA optimization," in *Computer-Aided Control System Design, 1994. Proceedings., IEEE/IFAC Joint Symposium On*, 1994, pp. 83-88.
- [14] Z. Michalewicz, Genetic Algorithms Data Structures= Evolution Programs. *Springer Science & Business Media*, 2013.
- [15] K. H. Lee, First Course on Fuzzy Theory and Applications. *Springer Science & Business Media*, 2006.
- [16] S. Chi and T. Cho, "Fuzzy logic based propagation limiting method for message routing in wireless sensor networks," *Computational Science and its Applications-ICCSA 2006*, pp. 58-67, 2006.
- [17] H. Y. Lee and T. H. Cho, "Optimized fuzzy adaptive filtering for ubiquitous sensor networks," *IEICE Trans. Commun.*, vol. 94, no. 6, pp. 1648-1656, 2011.

Crime Mapping in GIS by Using Hotspot

Syeda Ambreen Zahra¹

Abstract:

Police forces use hotspot mapping to provide a single out towards to resource allocation, ensuring police officers are posted to areas of high crime where their existence will have the most smash. Hotspot intelligence products are reliant on crime data sourced from police databases, and positional mistake in this data will have an impact on the accuracy of the hotspot maps make. The position of crime hotspots varies across both space and time. Crime mapping and analysis play an integral role in essentially advanced form of crime representation, visualization and respond satisfactorily to the problem of criminality. It also lets the analysts to figure out how crimes are spread evenly over the zone. GIS plays an effective role in mapping of crime. This paper puts on the diverse utilities of hotspot in GIS to recognize the crime in addition to encourage the advancement of investigation inclination strategy for policing. The functional approach in the present investigation for crime mapping can be successfully applied for improvement of user-interfaces stage for the advancement of safe city strategies.

Keywords: *Crime mapping; Spatial; GIS; Hotspot; Spatial Temporal analysis; Crime; Forecasts; Postional Errors; Predictive ability; Temporal information.*

1. Introduction

For the last few years a new worldwide socially order leads to the growing ratio on the criminal activities and raise the need to investigate latest methods to deal with information about criminality [1]. Crime mapping and spatial analysis of crimes are acknowledged as strong method for the learning and control of crimes because crime maps help one to investigate crime data and enhanced perceptible not only why a crime raises, but where it is taking place.

“Hurtful work or need against the masses as the State needs in similarity with stop yet which, above conviction, is culpable by means of fine, detainment, as well as death. No organization constitutes a crime unless it is pronounced pecan inside the lawful rules on the nation. Crime is unlawful attempt up to desire is denied by method for the law. Blame is habitually called an 'offense'. A few people put on shirts so much discourse 'it's exclusively unlawful if ye reach got' [2]”.

Straightforward maps, as show the areas the place violations or centralizations of wrongdoings hold came to fruition perform keep matured as per assist endorse watches in similarity with areas those are generally required. Approach producers inside police divisions may utilize more convoluted maps in congruity with inspect patterns among criminal movement, and maps may also demonstrate important between settling destructive cases.

For instance, analysts can likewise utilize maps as per better catch the looking examples on progressive hoodlums then as per speculate the place this guilty parties may live [28]. Utilizing maps so help people envision the geographic parts on wrongdoing, be that as it may, is presently not limited as per law requirement. Mapping is capable outfit particular records with respect to blame or criminal conduct in impersonation of lawmakers, the press, yet the general open.

Crime mapping answer three main subprograms within crime investigation [9].

¹ CS & IT, University of Lahore, Gujrat, Pakistan

Corresponding Email: msituol@gmail.com

- It uses visual and statistical analyses of the spatial conducting crimes and other types of actions.
- It allows analysts to associate spatial and non-spatial data.
- It furnishes maps that are helped to put across analysis results.

2. Crime Mapping Currently

Colleague effect neither provides careful, true and adequate matter not far from protect nor does it help in development goals and decision support. Spatial data analysis help one investigates crime data and enhanced perceptive not only why a crime rises, but where is taking place [3]. An acronym in light of Geographic records Systems that alludes after current portable workstation transcription up to desire catches, records, stores and examinations information regarding utilizations of earth's floor. It is additionally portrayed by method for paying for actualities alluding to as indicated by applications and in that place areas of floor surface sure as roadway, video show units exercises as much she happen, recovery or show of uncommon information, as pleasantly as, mapping [27].

It additionally involves geographic profiling the place areas are carefully entered by method for address, GIS thrived along the upward push concerning automated pc mechanical insurgency or has subsequently some separation measured as per remain completely phenomenal into settling many entangled social, monetary then politic inconveniences of humankind. Effectively, such has settled much injustice issues in the predominant world [26].

2.1. Getting Guilt to a Map

It is nonetheless workable in imitation of leading easy offense mapping through occupying pins between maps; alternatively crimes facts (both entire into entire or exclusively) contain a multiplicity on spatial-transient data [23].

Unless the records are mechanized then examined utilizing fitting programming, substantial tests also, clear procedures, so statistics intention remain usually inaccessible to both specialists also, professionals. The excellent programming arrangements are usually eluded in accordance with so geographic information frameworks, or GIS. GIS maintain spatial data of 3 essential ways: records are eking out away namely focuses, traces then polygons [22].

While spatial data are last as like focuses, traces and polygons, characteristic records are critical proviso the spatial records is in accordance with hold extra than shallow honor crimes records are mapped by a procedure called Geocoding [4].

Geographic statistics frameworks (GISs) improve PCs in conformity with speak according to and observe spatially associated wonder. All GISs bear twin functions:

- To show maps then geographic components, because example, obliqueness locations (focuses), streams (lines), yet assessment tracts (polygons)
- To utilizes a database supervisor so arranges and relates faith facts to specific information highlights.

A GIS use an advanced information database in conformity with interface spatial records in accordance with wise data. Several varieties over coordinating calculations possess a GIS in imitation of connect then preserve upon spatial connections amongst geographic yet enlightening data [25]. The potential according to interface or keep above spatial connections among statistics units characterizes a GIS [24]. The uncovering was undiminished close to handsome after goals

- To pick out warm spots as nicely as much using army because of specific sorts on crime.

- To help police in conformity with take strong measures kind of expanse regarding legion in area Inclined according to crime.
- To build over a methodological law because of wrongdoing mapping making use of GIS.

3. Geographical Information System and Crime Mapping

GIS play an essential role in crime mapping and analysis. The ability to contact and procedure information quickly, whereas displaying it in a spatial and visual means allows agencies to deal out assets rapidly and more successfully. The mainly dominant beat in law enforcement is information technology. Geographical information system is an information system that describes the objects with location [6].

A geographical information system convert physical elements of the real world such as roads, rivers, mountains, buildings into forms that can be visualized and analyzed, such as banking system, climate system, oceans, traffic, educational system police information about crimes,. GIS utilizing two sorts of information model vector and raster information. Vector deals with the discrete objects and raster deals the continuous objects [20]. Both vector and raster are not the same as every other After collection, edition and approving this data spatial analysis permits the assessment of these attributes and with the following space, it gives a geographical value to any geographical wonders. The usage of geographic data framework for wrongdoing mapping maps, envisions and analyzes crime hot spots along with other vogues and forms. It is a basic constitute of offence judgment and the protect deposit [7]. A GIS applies pair types of make to suggest objects and locations in the real world These are denoted to as polygon, point, line and image features.

The spatial data may be the location. GIS not only permits consolidation and spatial analysis of the data to discover, capture and

indict mistrust, but it also helps more positive measures in the course of helpful allowance of resources and better policy setting. In the next section of the paper we provide a framework of crime mapping that include background, methodologies and conclusion [18].

4. Background

Back the emergence of crime mapping technologies in the 1980's and 90's intelligence led policing has become a vital piece in the attack against crime for police forces all over the world. Analysts and police are expected to work hand in hand in the avoidance and distinguishing of crime and the proactive establishment of solutions [5].

Hotspot mapping is described by Chainey as one of the nearly accepted procedures used by crime analysts to exact police solutions [21]. In these times of austerity, with decreased numbers of officers it becomes even more major for police forces to have a select approximate, utilizing hotspot intelligence to make sure that police solutions are intensive on areas of high crime concur that where hotspot policing attempt are in place a clear trimming in crime is perceived, though the rationality of any hotspot mapping approximate is only ever as good as the source crime data used to create these maps in the first location [19].

5. Literature Review

Till our study for this proposal shows that despite the advancement in the technological field still outdated methods are being used for mapping and tracking the crime in the society. This practice of outdated methods produces a major gap between the response from police or action bodies and the criminals [9]. This happens due to the use of slow method (i.e. pin maps) to detect the crime location. The outdated method makes a huge waste of manpower and time resources, and produces frustration in the police department. Hence the latest technology like Geographical Information System (GIS) must be used to map and detect the crime location in order to make quick response to the crimes [8].

6. Methodology

After reviewing several papers, we have summarized different methodologies that help Authors to understand spatial analysis of crime.

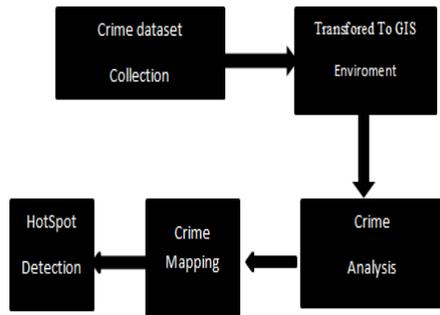


Figure1. Crime mapping by using Hotspot Detection

• Identifying Hotspots

Hotspot technique is basically used to identify areas where crime levels are high. The Hotspot analysis tool identifies spatial clusters of statistically significant high or low value attributes. Different methods for hotspot detection are as follows [11]:

1. Data Acquisition

Crime data from any GIS could be calculated in spatial and attribute or non-spatial data and there are diverse procedure for getting these data.

2. Spatial and Non Spatial Data Acquisition

In Data acquisition crime mapping and investigation establish GIS scenario can be done by using Interpolation Method. The data were derived in the form of latitude and longitudinal values of the given area and it was fed in to an excel file (Windows 2007). After his, the file was bring and change to GIS territory. In this paper Hotspot is used for crime mapping and analysis [12].

3. Interpolation by using Hotspot Technique

Interpolation methods are used to compute the unrevealed gap of attentive points by mentioning to higher gap data of adjacent points. Between many Hotspot techniques.

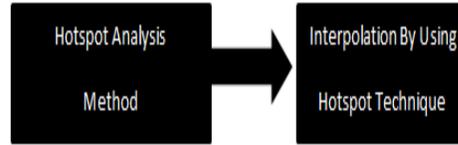


Figure 2

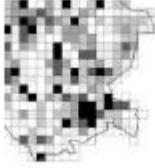
• Generating Hotspots

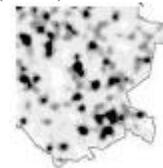
Hotspot mapping is established on the hypothesis that region of high crime will exhibit as a clustering of points in a spatial distribution. The elevation of clustering in the data will impact the predictive capability of the hotspots construct [10].

The crime reports of 2015 and 2016 were geocoded with the XY locations. Only the network constrained crime events were included in the analysis. House robberies, burglaries, crimes against persons, and other petty crimes are not part of this analysis. A total number of 2059 crime events in 2015 and 2016 were geocoded. The points were snapped to the lines by using the RTW tools for ArcGIS developed by Wilson in 2015 [17].

Table I. Overview of the four most common hotspot mapping Techniques.

Hotspot procedure	How it works?	Benefits
Spatial Ellipses	Uses STAC (Spatial and Temporal Analysis of CRIME) application to identify hotspot	Size and alignment of each hotspot is easily visualized. No reliance on defined geographical boundaries.

	areas and fit a standard deviation ellipse to each hotspot area.	
Grid Thematic Mapping 	Uniform grids are drawn over a study area and thematically shaded based on the number of crime points within each grid square	Equally sized grid squares means hotspot areas can be easily identified without risk of misinterpretation
Thematic Mapping of Geographic Areas 	Hotspot areas are based on defined administrative or political areas. Each area is thematically mapped based on the number of crimes occurring within them	Reflects areas and boundaries used by an organization. Thematic map produced is logical and easy to understand.
Kernel Density	Produces a	Visually effective.

Estimation (KDE) 	continuous surface through aggregation of point data within a specified search radius	Representative of spatial distribution of crime events. No reliance on defined geographical boundaries.
--	---	---

7. Conclusion

Crime mapping and analysis techniques are used to find crime hotspot locations. Forecast about crime is a tall order; we are not on final stage where we define specific events by a special offender at specific movement in the crime [13]. Using GIS into law enforcement has been an important dramatic for crime analyst and criminal justice researchers [28]. To keep crime analysis and decision-making, we need to recognize the complex (spatial) clustering (block) analysis. Hotspot provide crime analyst graphics representation of crime-related problems. Perceiving where and why crimes occur, can improve the struggle to fight for crime [15].

Using good management hotspot techniques we can reduce crime rates. We need to follow new technology in the 21st century to prohibit crime [20]. Eventually, Hotspot mapping and GIS support regional and complicated oriented policing GIS and Mapping can show the comprehensive correlation between the crime, the victim and the offenders. Some important facts of GIS and Crime mapping are: showing the probability and people changes, help in resource allocation, combining data from the government resources and community, providing effective communication tools [14].

Whatever approaches makes sense for you, applying and studying hotspot into law and defense is a twice successful choice [21]. As you advance, your own career should make important addition for social freedom and

order. Think of it as two benefits for one effort [16].

ACKNOWLEDGMENT

We authors acknowledge with thanks the assistance rendered by Dr. Javed Anjum Sheikh, University of Lahore, Gujrat Campus for providing crucial insight during the course of the research work which greatly improved the manuscript.

REFERENCES

- [1] Y. Bello et al., "Principal Component Analysis of Crime Victimization in Katsina Senatorial Zone," *International Journal of Science and Technology*, vol. 3, no. 4, pp. 192- 202, 2014.
- [2] P. Yar and J. Nasir, "GIS Based Spatial and Temporal Analysis of Crimes, a Case Study of Mardan City, Pakistan," *International Journal of Geosciences*, vol. 7, no. 19, pp. 325- 334, 2016.
- [3] T.F. Balogun et al., "Crime Mapping in Nigeria using GIS," *Journal of Geographic Information System*, vol. 6, no. 5, pp. 453- 466, 2014.
- [4] S.M. Ansari and Dr. K.V. Kale, "Methods of Crime Analysis using GIS," *International Journal of Scientific and Engineering Research*, vol. 5, no. 12, pp. 1330- 1336, 2014.
- [5] U.S. Usman, "The Role of Geographic Information System (GIS) in Effective Control of Terrorism in Nigeria," *International Journal of Economics, Commerce and Management*, vol. 3, no. 4, pp. 1- 9, 2015.
- [6] 2016. [Online]. Available: http://us.corwin.com/sites/default/files/ubinary/6244_Chapter_4_Boba_Final_PDF_3.pdf. [Accessed: 23- Nov- 2016].
- [7] Crimemapping info, 2016. [Online]. Available: <http://crimemapping.info/wp-content/uploads/2015/07/CMN3PDFv4.pdf>. [Accessed: 23- Nov- 2016].
- [8] Esri.com, 2016. [Online]. Available: <http://www.esri.com/software/arcgis/arcgisonline>. [Accessed: 11- oct- 2016].
- [9] "Crime Mapping and spatial Analysis," ITC.nl. [Online]. Available: http://www.itc.nl/library/papers_2003/m-sc/gfm/ahmadi.pdf. [Accessed: 27- Nov- 2016].
- [10] F. Fajemirokun et al., "A GIS Approach to Crime Mapping and Management in Nigeria: A Case Study of Victoria Island Lagos," Nigeria.
- [11] "Crime Mapping & Analysis News," Crime Mapping and Analysis News, 2016. [Online]. Available: <https://crimemapping.info/wp-content/uploads/2016/12/CMAN-Issue-5.pdf>. [Accessed: 17- Oct- 2016].
- [12] N. Levine, "Crime Mapping and the Crimestat Program," Wiley Online Library. [Online]. Available: <http://onlinelibrary.wiley.com/doi/10.1111/j.0016-7363.2005.00673.x/full>. [Accessed: 22- Nov- 2016].
- [13] A. Ahmed & R. Saliha, "Spatiotemporal pattern of crime using GIS approach in Dala L.G.A. Kano state, Nigeria," *American Journal of Engineering Research*, vol. 2, no. 3, 2013.
- [14] G.O. Igbaekemen & S.U. Umar, "A Purview into the historical Development of Terrorism in Nigeria," *Journal of Developing country Studies*, vol. 4, no. 14, 2014.
- [15] O. James, "GIS and Crime Management in Nigeria," *International Journal of Sociology, Management and Security Studies*, Maiden Edition, Kano, Nigeria, 2013.
- [16] F. Onu, "Corrupt Practices in Nigeria Society," *A Search For Causes and Remedies. IJMSSS*, Kano State, Nigeria, vol. 1, no. 1, 2014.
- [17] J. Shekwo, "Juvenile Delinquency in Mararaba, Karu L.G.A. of Nasarawa State, Nigeria," *International Journal of Sociology, Management & Security Studies*, Maiden Edition, Kano, Nigeria, 2013.

- [18] S.A. Yelwa & Y. Bello, "Complimenting GIS and Cluster Analysis in Assessing Property Crime in Katsina State, Nigeria," *American International Journal of Contemporary Research*, Vol. 2, no. 7, 2012.
- [19] S. Pires "Crime mapping and analyses news," *International Journal of Science and Technology*, vol. 4, no. 5, pp. 1-30 , 2012.
- [20] M.A.P chamikara., "GIS in crime analysis," *International Journal of Science and Technology*, vol. 3, no. 6, pp. 3 , 2014.
- [21] J. Bueermann., "Crime analysis," *Journal of Environment and Earth Science*, vol. 2, no. 3, pp. 1-6 , 2012.
- [22] M.brokmaan et al., "Crime Mapping and Analysis in the Dansoman Police Subdivision," *Journal of Environment and Earth Science*, vol. 4, no. 3, pp. 1-11, 2014.
- [23] C.D.J beaty., "GIS for Crime Analysis, Law Enforcement, GIS and Public Safety," *Journal of Environment and Earth Science*, vol. 4, no. 3, pp. 1-17, 2012.
- [24] T.Fransic et al., "Crime Mapping in Nigeria," scrip.org.[Online].Available: <http://www.scrip.org/journal/PaperInformation.aspx?PaperID=50296>. [Accessed: 20- Nov- 2016].
- [25] J.corso et al., "Toward Predictive Crime Analysis via Social Media, Big Data, and GIS," *Journal of Environment and Earth Science*, vol. 2, no. 3, pp. 1-6 , 2015.
- [26] S.Muhammad et al., " Mapping and Analysis of Crime in Aurangabad City using GIS," *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 2, no. 4, pp. 67-76 , 2014.
- [27] F.Wang., "Why Police and Policing need GIS," *Journal of Environment and Earth Science*, vol. 2, no. 4, pp. 67-76 , 2012.
- [28] T.Balogan et al., " Crime Mapping in Nigeria Using GIS," *Journal of Geographic Information System*, vol.6, no. 4, pp. 453-466 , 2014.

Evaluation of Android Malware Detectors

Hassan Rafiq¹, Muhammad Aleem¹, Muhammad Arshad Islam¹

Abstract:

Malware is an umbrella term used for viruses, worms, and Trojans. These days malware is becoming a great threat to the Android users. A malware detector which is commonly known as an antivirus or virus scanner avoids a malicious file to infiltrate into a system. With the increasing usage of smartphones, malware is also becoming powerful to penetrate the mobile devices. Traditional protection systems identify malware using signatures that can be manipulated by various techniques. In this research paper, it has been demonstrated that the most of the known commercial malware detectors cannot detect common code obfuscation techniques. Moreover, we have evaluated resource utilization (CPU, memory, and battery) consumed by several malware detectors.

Keywords: *Malware detectors; Antivirus; Android; Hardware performance counter.*

1. Introduction

A malware can penetrate into the host devices through the several ways. For example, a malware can integrate itself with a downloaded file downloaded, or via infected flash drives, or someone can intentionally insert a malicious file into a system. A malware developer can spread a malicious file via email or by attaching it to an application which apparently seems to be legitimate. Generally, malware can be classified on the basis of the propagation methods as discussed by McGraw et al. [1].

A malware can cause a severe damage to the infected devices, for example, it can compromise *confidentiality*, *integrity*, and *availability* of a system or network. Similarly, keylogger class malware can penetrate into a system to steal passwords and other sensitive information. Additionally, a particular type of malware commonly known as ransomware [2] encrypts the data and demand money for the data to be decrypted. Thus a malware can cause loss of important data and also cause huge financial loss to organizations and individuals.

Given the widespread emergence of Android malware, there is a crucial need to adequately moderate or protect against these threats. As indicated by the Intel Security/McAfee April 2017² patterns report; towards the end of the year 2016, there were more than 600 million malware variations altogether. There were approximately 15 million distinctive portable malware variations by the end of the year 2016. According to this report, nearly 08% of mobile users have been infected by some kind of smartphone-based malware. Thus, without an in-depth understanding of mobile malware, it is impractical to develop a reliable solution for the detection of mobile malware. In contrast to the existing mobile operating systems, Android is targeted mostly due to the open-source availability of this operating system [3].

A malware detector or antivirus identifies and scans a file using various mechanisms and checks whether the file is infected (malicious) or benign [4]. Generally, a malware detector executes in a passive mode (in the background) and scans a suspicious file. An

¹ Capital University of Science and Technology, Islamabad
Corresponding Email: aleem@cust.edu.pk

antivirus scans a file whenever a file is accessed or it performs a complete system scan on an explicit user request to scan every file in the system. Generally, a full system scan is applied and helpful when a user has installed an antivirus program (first time) and wants to ensure that there are no malicious programs in a system.

Similar to the personal computers, traditional approaches have been adopted to protect mobile devices too from malware threat. Mostly, malware detectors rely on the virus definitions to detect malware. These virus definitions are updated regularly i.e., every day or more often. Virus definitions mostly consist of signatures of the known malware families and variants. Malware detectors have to continually keep up-to-date with the latest malware definitions to be effective for malware detection. Antivirus tools employ a variety of tools to disassemble malware for analysis. Malware detectors also employ heuristics [4] which make a malware detector more capable to identify new malware even without the up-to-date virus definitions.

In this paper, we have highlighted a potential problem that the most of the commercial malware detectors are unable to detect obfuscated malware samples. With code obfuscation, a developer can hide the original algorithm or the logic of the code [3]. We have experimented using various types of code obfuscation techniques (as listed in Table 4) to benchmark which malware detectors are still able to identify a malicious code obfuscated within a legitimate application. Additionally, one of the key aspects of mobile devices is energy conservation. Therefore, the malware detectors are evaluated on the basis of resource consumption reported by the key performance counters including battery consumption. Our research aims are to benchmark the effectiveness of malware detectors against the obfuscated malware. Following are some of the contributions of this work:

- Using several types of code obfuscation techniques to test Android malware detectors;
- Benchmarking android malware detectors based on their malware detection capability;
- Profiling and analysis of Android malware detectors based on resource usages such as CPU, memory, and energy.

The structure of the rest of the paper is as follows. Section 2 discusses the related previous research works. In Section 3, we present the proposed methodology for benchmarking Android malware detectors. Section 4 presents the experimental results and discussion. Section 5 concludes the paper.

2. Background and Related Work

Android applications are developed in Java. Java source code is packaged into an Application (Apk) file which executes on the Android devices [5]. We use dex2jar [6] and apktool [7] to convert the android applications into the source code. After de-compilation into code and resource files, the Apks can be analyzed. In this paper, we de-compile known malware samples and make changes to their code without modifying the applications' functionality.

2.1. Code Obfuscation Techniques

Code obfuscation [8] is mainly done to hide the logic of the code so that the code could not be understood after reverse engineering. Code obfuscation changes the size and content of the Apk file; however, the main logic of the code is not modified. Code obfuscation does not have any impact on the semantics of a code. There are many code obfuscation techniques which can be applied to generate various code versions with the same semantics.

In one of the recent work, Zheng et al. [9] evaluated malware detection capabilities of malware detectors by applying code transformations. The authors developed

different test cases of malware samples by using several transformations and then evaluated using virus total [10] platform.

Authors employ artificial code diversity [11] as a code obfuscation method to evaluate the malware detection platform i.e., virus total. The authors prepared malware samples using a tool named *ADAM*. This tool was developed by the authors and employed for the code obfuscation. As compared to this work, we manually applied several obfuscation techniques after reverse engineering the malicious Apk file. Moreover, we perform testing on well-known commercial malware detectors.

Christodorescu and Jha [12] tested desktop malware detectors in the similar manner as we perform in this study. The results of their experiments show that the most of the malware detectors are unable to detect malware samples. Moreover, we experiment using six malware detectors as compared to three tested by the authors in [12].

Collberg et al. [8] have discussed different

kinds of code obfuscation techniques. They presented working and architecture of Java code obfuscating tool named as *Kava*. We use some of the mentioned code obfuscation techniques presented by the authors in [8].

Christodorescu et al. [13] have proposed a technique that suggests that the obfuscated malware samples can be detected. However, this detection is limited to detection of only garbage and re-ordered code. In this work, we use six code obfuscation techniques and their combinations as compared to only three employed by the authors to benchmarks malware detectors.

Protsenko et al. [14] have proposed a tool named as *Pandora* using can be used to apply obfuscation. After that, the obfuscated code can be tested using a malware detection tool such virus total. In contrast, we perform benchmarking of malware detectors using commonly used code obfuscation techniques and six most used malware detectors. In Table 1, a brief summary of the related work is shown.

TABLE 1. Related work summary.

Reference and Methodology	Strengths	Weaknesses
- Semantics persevering obfuscation techniques are applied.	-Obfuscated samples bypass malware detectors.	-Only three malware detectors are tested.
- [6], Different levels of obfuscation are used. -Each level consists of different combinations of code obfuscation techniques.	-Checks software plagiarism based on the proposed technique	-Testing of malware samples is performed on virus total only.
- [10], Variants of a single malware sample are prepared -Each sample is tested using malware detector “virus total”	-Malware samples are automatically prepared using a tool ADAM.	-Testing of malware samples is performed on virus total only.
-[14], Proposed a semantic-aware malware detection technique.	-Can detect a malware sample in which code obfuscation is applied	-Can detect malware based on only garbage insertion, code reordering, and register renaming based
-Proposed a mechanism to detect malicious files -Detects malicious files based on their behavior on the network.	-Obfuscated malware samples can also be detected	-Only applicable for malware which access network excessively

3. Methodology

The detection capability of malware detectors is tested by using obfuscated malware samples which are prepared by performing several different steps as shown in Figure 1. Only those malware samples are taken for code obfuscation which are detected as malware in the original form. (i.e. Before applying code obfuscation).

We prepare six different malware samples from a single malware by applying different code obfuscation techniques. The employed six code obfuscation techniques are listed below:

1. Variable Renaming [11]
2. Package Renaming [13]
3. Method Renaming [9]
4. Garbage Insertion [6]
5. Rebuilding [14]
6. Call Indirection [13]

1) **Variable Renaming:** All the variable names are modified in the context of variable renaming. Figure 1 shows an example code obfuscation using variable renaming.

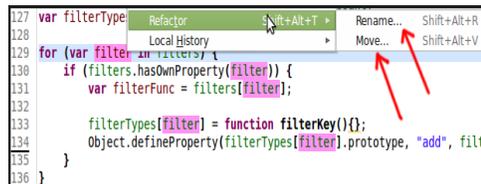


Fig. 1. Variable renaming.

- 2) **Package Renaming:** is related to changing package names of a given apk using the *Android Manifest* file.
- 3) **Method Renaming:** similarly, in method renaming names of all the method is changed.
- 4) **Garbage Insertion:** Whereas in garbage insertion, a garbage code is inserted that does not change the semantics of the application.

Listing 1: Indirect function call and garbage code insertion.

```

1 void display()
2 {
3   cout<<"hello world";
4 }
5 void show()
6 {
7   display();
8 }
9
10 void main( )
11 {
12  display(); //Direct call
13  //function
14  Show(); //Indirect call
15  //function
16  while(0)
17  {
18    cout<<"garbage";
19  }
20 }

```

In Listing 1, a *while loop* is shown with a false condition. The execution control never enters such loop and the enclosed code will not be executed. Such kind of code is referred as garbage code and can be inserted by the malware writers to create code level dissimilarity in malware applications.

- 5) **Rebuilding:** Another effective technique that can be used to test the malware detection capability of an antivirus or malware detector is *code rebuilding*. When rebuilding a code, the Apk is first decompiled and then is recompiled without making any changes in its resources and manifest file. Rebuilding process does not change the content of the Apk; however, it generally changes the byte order [9] and the hash value of the application. In most of the malware detectors, the detection algorithms mainly rely on the hash signatures of the files under investigation. Therefore, malware writers exploit this fact to doge the malware analysis tools.
- 6) **Call indirection:** is another effective technique in which the original method calls are re-programmed and shifted in some dummy methods to make indirect function

calls. Listing 1 at line 14 shows an example of *call indirection*.

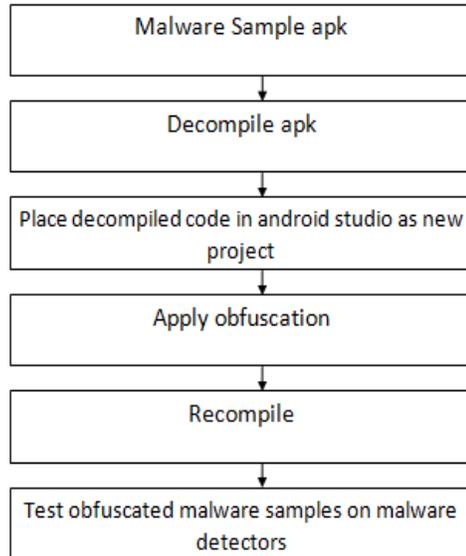


Fig. 2. Proposed methodology.

Figure 2 shows the proposed methodology used to benchmark malware detectors. The first step shows that a sample malware Apk is taken. We use malware application dataset available at [15]. The malware sample may belong to any known malware family. We choose only those malware samples which could be detected by the employed six malware detectors.

The second step of the methodology is based on decompiling Apks using dex2jar [6] and apktool [7] into Java code. In the third step, a new Android project is created based on the decompiled Java code and XML design files.

In the fourth step, obfuscation is applied to the decompiled code. To obfuscate the code, we employ the six obfuscation techniques and their combinations. The output of each obfuscated method is a new version of the Apk; for example, after changing names of all the variables the code is recompiled the version of the Apk is saved separately. To insert garbage-code, a redundant non-executable code is inserted (as shown in Listing 1) and is recompiled to generate the

Apk. Similarly, method calls indirection is used to invoke methods via some other indirect method as shown in lines 5-8 in Listing 1. In addition to the six obfuscation techniques, several other combinations (shown in Table 4) are used to generate several versions of the Apks.

4. Results and Discussion

4.1. Dataset and Experimental Setup

The experimentations were performed using an Android system with following specifications, i.e., CPU 1.3GHZ, quad-core, 01GBs of main memory, battery 200 mAH and Android version 4.2 (jelly beans). Table 2 shows the names of the Android malware detectors which have been tested.

TABLE 2. Malware detectors evaluated.

Product name	Total downloads (millions)
Norton Mobile Security	5M-10M
AntiVirus Free	50M-100M
ESET mobile security	500K-1M
Dr Web	10M-50M
Lookout mobile security	10M-50M
Zoner Antivirus	1M-5M

TABLE 3. Malware samples used for testing.

Malware	Details
Love Trap	A trojan that sends SMS
DroidDream	Creates spoof version of the original application
FakePlayer	Advertises unwanted products
Bgserv	Fake mobile cleanup tool
Basebridge	Performs harmful actions without user's knowledge
Plankton	Sends host's information to a remote server
Geinim-A	Corrupts the applications
LuckyCat	Opens backdoor in application to steal information
HippoSMS	Sends SMS to a hard-coded number
NickySpy	Sends host's information to a remote server

Table 3 shows the names and functionality of the malware samples which are used to prepare the test cases to evaluate the malware detectors shown in Table 2.

All the malware samples shown in Table 3 are detected as malicious in their original form (i.e., before obfuscation is applied) by all the malware detectors shown in Table 2. Some of the malware detectors have been omitted from Table 2 because they were unable to detect the malware samples as malicious which are shown in Table 3. All the malware detectors are directly downloaded from the official Android application market i.e., Google Play. Table 4 shows the list of obfuscation techniques that have been used in this paper to evaluate malware detectors.

TABLE 4. Labels of code obfuscation techniques.

Labels	Technique
VR	Variable Renaming
MR	Method Renaming
REB	Rebuilding
GCI	Garbage code insertion
RP	Package renaming
CI	Call indirection

4.2. Results

Table 5 shows the minimal combinations of obfuscation techniques required to evade a malware detector. For example, *LoveTrap* requires variable renaming, method renaming, and package renaming to evade Norton antivirus, Antivirus free, ESET and Lookout. *Love Trap* remains undetected by Dr. Web if package renaming and call indirection is applied, whereas *Zoner* cannot detect *LoveTrap* if simple rebuilding is applied to it.

Similarly, when we consider *DroidDream* malware sample then the results shown in Table 5 highlight that the Norton and the ESET cannot detect *DroidDream* sample for the combination of package renaming and rebuilding obfuscations. Whereas, in case of Dr. Web malware detector, the *Lookout*, *Zoner*, and the *DroidDream* samples

go undetected (with simple application rebuilding obfuscations).

In case of *Bgserv* malware sample, the results of Table 5 show that the Norton malware detector is evaded by the obfuscation combination of package renaming, variable renaming, and method renaming. The *AntiVirus free* is evaded by the obfuscation combination of package renaming and call indirection. The malware sample *Bgserv* could not be detected as malicious by the ESET and *Lookouttools* for the obfuscation combinations based on package, variable, and method renaming. The malware detector Dr. Web also could not detect *Bgserv* malware sample based on call indirection obfuscation. The malware detector *Zoner* could not detect *Bgserv* as malicious even when a simple rebuilding was applied to it. The *Hippo SMS* malware evaded malware detection capability of *Antivirus Free*, *ESET*, and *Dr. Web* when a combination of package renaming and rebuilding was applied. The *hippo SMS* evaded *Lookout* and *Zoner* malware detectors when the malware sample was simply rebuilt.

Keeping in view the results of Table 5, we may conclude that the Norton antivirus is a hard nut to crack because it can only be evaded if complex obfuscation is applied to a malware sample i.e., a combination of variable renaming, method renaming, and package renaming. On the other hand, the *Zoner* malware detector proves to be the weakest among the employed anti-malware because it can be evaded by simply re-building a malware sample. The rest of the malware detectors (as shown in Table 5) are not resilient to several combinations of code obfuscation techniques. Most of the malware detectors are able to detect the re-build samples; however, they are unable to detect malware samples when several obfuscations are used collectively.

Figure 3 shows the malware detection results for different malware detectors against the employed obfuscation techniques. In Figure 3, the Y-axis shows the tested malware detectors and X-axis shows the percentage of

samples evaded the employed malware detectors. Figure 3 presents the results of the malware sample *Hippo SMS* and its versions based on code obfuscation. For Norton antivirus, the results show that most of the code obfuscations have been detected; however, the combination of *Package Renaming* (RP), *Variable Renaming* (VR), and *Method Renaming* (MR) obfuscation techniques resulted in 70% undetected cases. For the combination of *package renaming* and *rebuilding* results in only 20% of un-detected cases for the Norton antivirus. In our experiments, we observed that the *variable renaming*, *method renaming*, *package renaming*, *call indirection*, and simple rebuilding are easily detectable using the Norton antivirus. Moreover, Figure 3 shows the detection results of other antiviruses for the employed code obfuscation methods. As shown in Figure 3, simple code rebuilding is detected by most of the antiviruses except *Dr web* (30% samples undetected) and *lookout* (10% samples undetected). The combined obfuscation based on *package renaming* and *call indirection* also show a large percentage of un-detectable malware samples. The results show that the most stealth obfuscation samples were based on the combination of *package*, *method*, and *variable* renaming. Similarly, a higher evading result was shown for the code obfuscations based on simple package renaming combined with call indirect.

Next, we perform resource consumption analysis for the employed 06 android malware detectors. Table 6 shows the results obtained using the benchmarking tool Mobibench [16]. Table 6 presents the resource consumption chart for 06 malware considering the CPU, memory, battery, and storage requirements. Mobibench employs Android APIs to calculate memory and processor usage. To calculate battery consumed by a malware detector, the Mobibench requires an Android device to run in a *clean state* (i.e., no other application being executed at that time of instance). Mobibench records the battery level

of the device before starting the malware detector and again record the battery level after the malware detector finishes its task (of screening). The battery consumed is shown in units milli-ampere-hour (mAH) as shown in Table 6.

Table 6 shows that *Dr Web* consumes 16% CPU, 56% RAM or memory, 0.91 mAH battery, and 7.13 MBs size on disk. Similarly, the performance analysis of other malware detectors is shown in Table 6. These results show that the Norton antivirus is the highest resource consuming malware detector whereas the zoner malware detector consumes the least device resources as compared to other malware detectors.

5. Conclusion

The experiments performed in this research show that there are serious shortcomings in the commercially available malware detectors (against the obfuscated malware). To demonstrate these, we employ several malware detectors and tested those using many combinations of code obfuscations. Most of the time, an obfuscated malware is undetectable. However, a few Android malware detectors (such as Norton, antivirus free, etc.) are able to detect malware obfuscated using multiple techniques. The results clearly show that well known commercial malware detectors are not resilient to common code obfuscation techniques. In addition to this, it has also been observed that the malware detectors which have good detection rate also consume more device resources especially battery and storage space. In future, we intend to research the mechanism using which a malware detector should be able to detect the obfuscation applied to the original malware sample; hence, improving overall malware detection rate.

TABLE 5. Evaluation summary.

	Love Trap	Droid Dream	Fake Player	Bgserv	Base Bride	Plankton	Geinim-A	Lucky Cat	HippoSMS	NickySp.y.B
Norton	RP+VR+MR	RP+REB	RP+VR	RP+VR+MR	VR+MR	RP+VR+MR	RP+VR+MR	VR+MR	RP+VR+MR	RP+VR+MR
Antivirus Free	RP+VR+MR	VR+MR	RP+VR+MR	RP+CI	RP+CI	RP+VR+MR	VR+MR	RP+REB	RP+CI	RP+VR+MR
ESET	RP+VR+MR	RP+REB	CI	RP+VR+MR	RP+VR+MR	RP+CI	RP+REB	RP+REB	RP+CI	RP+CI
Dr. Web	RP+CI	REB	RP+REB	CI	CI	CI	RP+REB	RP+REB	RP+CI	RP+REB
Lookout	RP+VR+MR	REB	RP+VR	RP+REB	RP+CI	REB	REB	REB	REB	RP+REB
Zoner	REB	REB	REB	REB	REB	REB	REB	REB	REB	REB

TABLE 6. Resources consumption analysis.

Antivirus	CPU(%)	RAM(%)	Battery (mAH)	Storage Size (MB)
Dr Web	16	56	0.91	7.13
AntiVirus Free	22	65	0.84	4.7
Lookout	18	69	0.9	9.05
Norton	30	60	1	17.15
ESET	21	64	0.98	7.93
Zoner	19	56	0.8	1.56

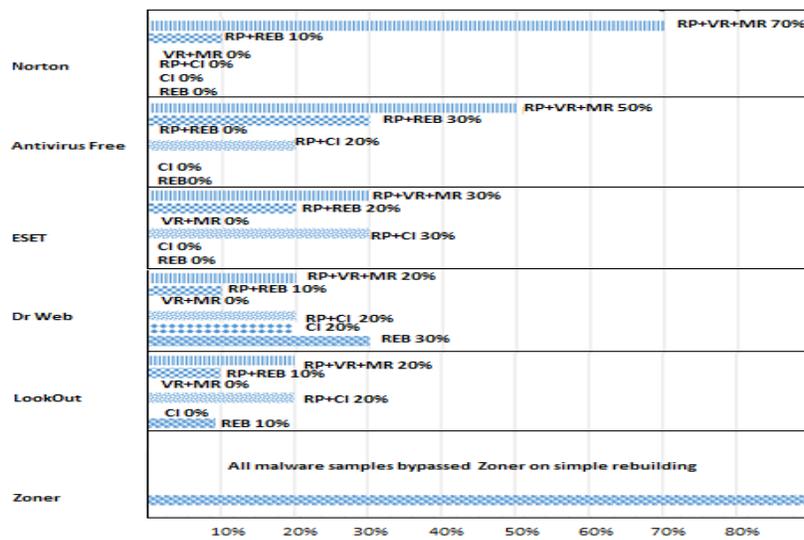


Fig. 3. Experimentation results.

REFERENCES

- [1] G. McGraw and G. Morrisett, "Attacking malicious Code: report to the InfoSec research council," *IEEE Software Magazine*, vol. 17, no. 5, Sep.-Oct., 2000.
- [2] S. Aurangzeb, M. Aleem, M. A. Iqbal, and M. A. Islam, "Ransomware: A Survey and Trends," *Journal of Information Assurance & Security*, vol. 6, no. 2, 2017.
- [3] M. P. Dalla and F. Maggi, "Testing android malware detectors against code obfuscation: a systematization of knowledge and unified methodology," *Journal of Computer Virology and Hacking Techniques*, vol. 13, no. 3, 2017.
- [4] "How antivirus works.," [Online]. Available: <https://goo.gl/4HxMu1>. [Accessed 23 7 2017].
- [5] "Android Developers.," [Online]. Available: <https://goo.gl/q9sLWI..> [Accessed 22 5 2017].
- [6] "Dex2jar," [Online]. Available: [http://code.google.com/p/dex2jar/..](http://code.google.com/p/dex2jar/) [Accessed 22 5 2017].
- [7] "Apktool.," [Online]. Available: [http://code.google.com/p/apktool/..](http://code.google.com/p/apktool/) [Accessed 22 5 2017].
- [8] C. Collberg, C. Thomborson, and D. Low, "A Taxonomy of Obfuscating Transformations," *Department of Computer Science, The University of Auckland, New Zealand*, 1997.
- [9] M. Zheng, P. P. Lee, and J. C. Lui, "ADAM: an automatic and extensible platform to stress test android," in *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, Springer Berlin/Heidelberg.
- [10] "VirusTotal," [Online]. Available: <https://goo.gl/DITruF>. [Accessed 22 5 2017].
- [11] J. Nagra, C. Thomborson, and C. Collberg, "A functional taxonomy for software watermarking," *Australian Computer Science Communications*, vol. 24, no. 1, pp. 177-186, 2002.
- [12] M. Christodorescu and S. Jha, "Testing malware detectors.," *ACM SIGSOFT Software Engineering Notes*, vol. 29, no. 4, pp. 34-44, 2004.
- [13] M. Christodorescu, S. Jha, S. Seshia, D. Song, and R. E. Bryant, "Semantics-aware malware detection," *IEEE symposium on Security and Privacy*, 2005.
- [14] M. Protsenko and T. Muller, "Pandora applies non-deterministic obfuscation randomly to android.," in *"The Americas" (MALWARE)*, 2013.
- [15] "Contagio minidump," [Online]. Available: <https://tinyurl.com/6b6v7jp>. [Accessed 27 5 2017].
- [16] A. Zaman and Z. Imtiaz, "MobiBench," *Capital University of Science and Technology, Islamabad.*, 2016.

Challenges of Computer Science and IT in Teaching-Learning in Saudi Arabia

Hafiz Abid Mahmood Malik¹, Faiza Abid², R. Kalaichelvi¹, Zeeshan Bhatti³

Abstract:

Personal Computers (PCs) have invaded all ranges of civilization and there is currently a reasonable connection among technology, development and economic persistence. For a couple of decades in teaching-learning field, computer science curriculum development has been a core issue. In different eras different strategies have been adopted to improve the teaching-learning process. Integrated Curriculum Techniques (ICT) can play a positive role in all subjects. Specifically, strengthen the integration among all subjects can prove the good results in computer science and English language. In the area of computer science and information technology, communication in English language is considered a big barrier in understanding in Saudi Arabia. In this research we discuss and suggest the different aspects that can upsurge the overall performance and structure of the education system, particularly in Saudi Arabia. It is an effort to investigate the nature of hindrances and challenges faced at Saudi academies while implementing an IT based learning approach. To walk with the level of international universities and to play the role in modern scientific research, the use of the English language can improve the vision of computer science and information technology. Teacher-student communication in English language is a very important factor. This small step-forward can develop the confidence in students that consequently leads to the betterment of whole education system. Besides, the availability of up-to-date and accurate software and hardware installation is very important to cope with the challenges of the era. World is in your hands, if you know the proper use of technology. In this paper, Makerspace techniques have also been suggested that can play a significant role in teaching-learning that relates with ICT.

Keywords: *Computer science; Information technology; Makerspace; Education; Challenges; Curriculum; English language skills.*

1. Introduction

The Ministry of Higher Education (HEC) in Saudi Arabia has encouraged the implications of Information Technology for teaching-learning among the students and teachers. Computer Science (CS) and Information Technology (IT) as speculative disciplines that deliver knowledge and proficiency substance for all hi-tech developments. Contrasting to other more stagnant disciplines, CS and IT are continuously being reformed. Innovations and new technologies continue to increase our

indulgent of what computer scientists can do and how much our lives can be relaxed [1]. Knowledge of CS and IT are currently much vital to today's scholars as any of the customary sciences. Computers have subverted all extents of civilization and there is now a clear link among technology, revolution and economic existence. Deficiency of general curriculum standards, consistent and cogent teacher accreditation requirements remain to hamper the ability to ensure that students are effectively prepared to compete in this progressively high-tech world. This is the era

¹AMA International University Bahrain

² King Khalid University, Saudi Arabia

³ University of Sindh

Corresponding Email: hamalik@amau.edu.bh

of science and technology. New innovations have changed the trend of the education now, and without a PC no one can complete his research, as you need internet all the time for the research and originations [2]. To be in touch with the new trends and global issues in all fields of education, everyone is requiring fundamental knowledge of computer and internet, at least. On the other hand, CS extents a wide range of computing activities from academic grounds to automation, computer visualization, smart and intelligent systems, nanotechnology and bioinformatics. It is a study of computers, hardware, software designs, their applications, and their influence on a society. In this study, different obstacles in the way of CS and IT teaching are discussed are some remedies are given to avoid those obstacles. Main locale of the study is Saudi Arabia. The study has discussed the ways to improve the teaching-learning abilities. Students should be prepared to compete in the increasingly technological world and they should be familiar with practical implementation in the real world problem.

2. Literature Review

From a survey report (Sloan Consortium Survey) about online education in US, in year 2011, Rebecca stated that online enrolments evaluation is ten times higher than traditional mode [3]. Furthermore, computer education is an essential part of the online education.

Due to electronically support education and training pedagogy for student hub and collective learning has turned into familiar. E-learning is a totally new learning platform for students and teachers, therefore, computer skills are involving for its implementation. Due to evolutions in Information Communication Technology (ICT), students study or learn without schools' places; therefore, teachers' responsibilities and students' learning practices are also changing [4].

Knowledge and skills are sources of success and needed in current education system (primary schools, high schools and

higher secondary schools) all over the world. According to R. Sims, e-learning model is new practice of learning and activity which improved and transformed traditional style of learning in the form of well-organized, effective and attractive new technology models [5].

The government of Pakistan in the province of Khyber Pakhtunkhwa, initiated IT Labs and laptop schemes, the KP pilot project distributed 2800 tablets to teachers but e-learning concept, understanding and policy for the learning in KP are still remaining problems [6]. Government only just focused on hardware rather than utilization of ICT for education.

Learners can get opportunities and enhance their knowledge from blended learning which combines face-to-face and online mode of instruction [7]. Below some philosophies of CS and IT in teaching learning are discussed.

3. Some philosophies of Computer Science and Information Technology in teaching-learning

Below are some important philosophies of computer science and information technology that should be considered in teaching-learning environment:

- Scholars should gain an extensive synopsis of the field to build a broad image of CS and IT as a discipline.
- Students should comprehend not only the theoretical reinforcements of these disciplines but also learn how that theory affects practice.
- CS and IT education should emphasis on problem unraveling and algorithmic philosophy.
- Concepts should be educated liberated of explicit applications and software design.
- Scholars should be trained what will be anticipated from them in "real world" precisely, what is truly required to compose and retain computer programs

and large software systems. CS and IT should be educated using real-world problems and applications instead of specific academic tools.

CS and IT education should comprise integrative and interdisciplinary acquaintance.

4. Methodology and Challenges

Although marvelous efforts have been exercised to improve the teaching-learning process of computer science and information technology but these could not produce the required outcomes. Below are some challenges that should be addressed while teaching-learning and these techniques can be utilized to achieve the above-mentioned objectives.

4.1. Curriculum development

It is very important to develop an up-to-date curriculum of computer science and IT which fulfills the recent requirements of the subjects. An advanced curriculum always attracts the sensible students who desired to walk with the world. A country should have to design and implement general CS academic programs in order to better formulate their students for the progressively reasonable universal economy and current trends [8].

4.2. Integrated curriculum

Integration is the association of teaching substance to interconnect or unify subjects frequently taught in distinct academic courses or departments. In 1980s and 1990s, "curriculum integration" was stated as interdisciplinary, multidisciplinary and trans-disciplinary curriculum designs. Pioneering educationalists concerned with improving student accomplishment are seeking ways to create rigorous, relevant, and engaging syllabus [9]. An incorporated and integrated curriculum work regarding association, it may be in real life or crossways the disciplines, regarding abilities or about understanding and knowledge [10] [11]. It rages subject varieties, capabilities and real life familiarity together to make a more accomplishing and perceptible

learning environment for learners. Another motive of an incorporated syllabus is that it compromises more recurrence of knowledge than to teach subjects in segregation [12]-[13]. So, there must be integration among all subjects regarding CS and IT.

4.3. Language skill

Communication is a significant part of teaching-learning process. Though we can communicate in any language to teach the students (e.g Arabic in Saudi Arabia) but we must see the international trends to teach on international standards. English is an international standard language that is mostly being used in all over the world's best institutions and mainly in research. Therefore, to cope with recent trends and research in teaching-learning environment communication must be in English language.

4.4. Computer labs

Up-to-date hardware and software equipped computer labs are necessary to accomplish and teach the subject of computer science and information technology. Proper computer networking is required to better control the computer systems as well as students. Furthermore, e-learning labs should be facilitated separately.

4.5. Interactive learning

It is a more hands-on, real-world process of relaying information in classrooms. Interactive instructional techniques discourse the need for pupils to be active in the learning practice and to interact with others. Scholars can reinforce their knowledge about coaching by interacting with you and with their colleagues. Interactive instructional stratagems provide chances for students to reinforce their observational skills, listening skills, communication skills, and interpersonal skills.

4.6. Deficiency of focusing on skill development

It has been observed that students do not emphasis on developing skills rather they

focus on rote learning. In general, most of the students excel when examinations emphasis mainly on memorization and rote learning. Besides this, they do not do well when assessments / papers contain critical thinking questions, creativity, tricky methodology, or logical problem solving. Students should focus on skill development as well instead of rote learning.

4.7. Role of makerspaces in teaching-learning activities

Makerspaces is the idea to provide separate places for teachers, students and researchers where they can share technology, ideas and knowledge to produce innovations. Through these collaborated work places many people can get new ideas and benefits in term of removing their confusions and to get more motivations towards their advance projects. Most of the time the quality of work is better in team work or in other words, more brains produce better ideas. So, Makerspaces are the good idea for the quality work. In every institution particularly in research oriented institutions there should be a Makerspace for the teaching-learning environment. Furthermore, it somehow cut the cost of research and technology equipment compare to individuals. Makerspace idea can be facilitated and implemented in form of separate e-learning labs.

4.8. Creating awareness of blended learning methods

Blended teaching is more constructive method than purely face-to-face classes. Blended instruction [14, 18] is a mixture of digital instruction and one-on-one face instruction in a traditional classroom environment. By combining information technology into class projects, students will be able to have better understanding of course material. Eventually, the students will be motivated to learn computer science and Information technology subjects and they implement the practical knowledge of computer science in to other subjects. Additionally, weaker students can be given

tutorial classes through online at their convenient time. Hence, the level of achievements is higher and effective in blended learning than face-to-face learning.

4.9. E-learning

E-learning is an interactive and flexible means of learning or getting information. Through E-learning, learning and teaching can happen at anytime and anywhere. The components of e-learning include text, audio, video and animation. Using E-Learning apps, students' motivation towards learning subjects can be boosted. Also, it allows students to enhance diversity, critical thinking for innovative problem solving. In other way, practice of E-learning makes faculty members familiar with current Internet trends and how they can be applied in the lessons. At the same time, it supports faculty members to generate their own courses and activities on the web-based platform.

Some of the E-learning tools: TalentLMS, Lessonly, Digital Chalk, Moodle, Socrative, Kahoot, TED-ED, Instructables, Hopscotch, Slack and uBoard digital interactive whiteboard [15-16]. By using the e-learning tools, the following activities can be incorporated through online.

- Problem – Solving Activities;
- Project – Solving Activities;
- Action – Based Strategies;
- Brainstorming;
- Questionnaires;
- Quizzes;
- Announcements;
- Forums.

4.10. Web-based education

The public in Saudi Arabia has been using internet since 1999. In this internet era 60% of the Saudi population are adapting emerging new technologies. Most of the Saudi universities use the traditional lecture based

classroom teaching method with few programs having distance learning. In distance learning programs, they use web-based instruction [14, 18]. The web based instruction facilitates online teaching and learning methods. Nevertheless, only a limited number of faculties using online teaching methods, those who have adequate skills in e-learning. In order to increase the e-learning techniques in Saudi based universities, proper training should be provided to all faculties. Besides, female faculties have negative perspective using internet as it has immoral content. This perception lessens the adoption of web-based education. Furthermore, the faculty members with inadequate knowledge of internet usage find difficulty in adopting web-based education. To discourse these issues, Saudi based universities should provide proper awareness, support services and adequate training to both faculty members and the students. These factors improve the confidence in using web based technology either in delivery approach or in learning environment.

4.11. National E-learning and Distance Learning Centre (NELC)

Universities shall provide digital educational system by using National E-learning and Distance Learning Centre (NELC) [14] established by the Ministry of Higher Education, Saudi Arabia. To implement e-learning in Saudi universities, NELC offers practical tools such as multimedia resources to empower faculties of universities to implement online teaching. NELC established “Jusur E- Learning Management System” (LMS) [17] to support teaching and learning system in Saudi universities. Jusur builds E-learning culture among the faculty members and the students. However, student discipline, responsiveness, and training of the LMS are the challenges in implementing Jusur. In addition, adequate training should be given to faculty members to use Jusur as their teaching strategy. Learning Management System (LMS) is shown in figure 1.

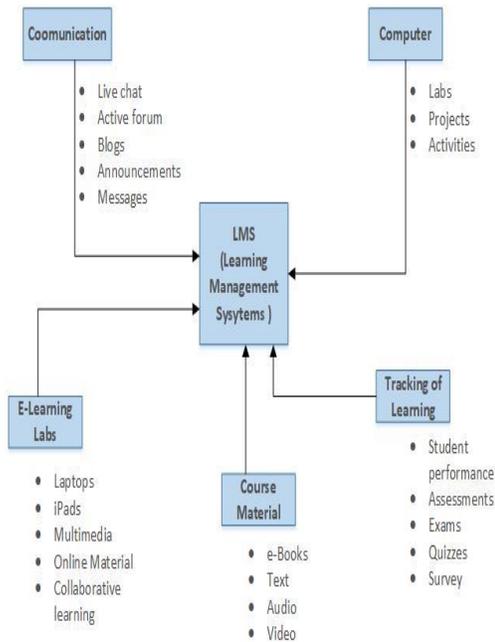


Fig. 1. Learning Management System.

4.12. Recommendations

- Academic staff should be only well qualified teachers
- An up-to-date curriculum of computer science and IT should be implemented which fulfills the recent requirements of the subjects.
- Instead of teacher-centered activities there should be learner-centered environment.
- Integrated curriculum should be practiced for better teaching-learning.
- Teacher-student communication must be in English language.
- Using teaching techniques that integrate language skills rather than teaching them discretely.
- Examinations should be regarded as a teaching and learning device rather than being just a testing device.

- Web-based education system can improve the quality of teaching-learning environment.
- E-learning technique can be helpful to enhance the educational standard.
- National E-learning and Distance Learning Centre (NELC) should be considered.
- National Center of e-learning and digital learning (NCeL) may be useful for improvements.
- Students can be categorized according to their learning skills.

5. Results

If authorities take the above suggestions into consideration then very positive outcomes are expected. Curriculum will be improved and there would be a proper integration among all subjects as well. Language skills of the students will be enhanced. Computer labs will be highly equipped and properly managed and utilized. In addition to students' development, teachers will also be improving their skills and overall teaching-learning environment will be positively affected.

6. Conclusion

World has become global village due to the advancements in technology. Computer science and information technology are the main pillars in these achievements. So, in this era, any country which does not get benefits from the science and technology will be backward. In the field of CS and IT teaching-learning is very important factor. Here, we have discussed some teaching challenges in the form of objectives and their remedies are suggested in the form of methods. Through the integrated curriculum techniques, we can upraise the quality of education. Further, by permanent communication in English language, student's confidence and his/her understanding caliber towards research and innovation can be enhanced. That would definitely be helpful in all subjects as well. By

that we can prepare the student to compete with the challenges in the upcoming technologies. Our main focus is student to whom we aimed to give quality education and knowledge. So that he/she may be a part of a healthy society and could play his/her role in the betterment of the country and humanity. Furthermore, Makerspaces can be positive step towards research and innovations.

REFERENCES

- [1] Stephenson C, Gal-Ezer J, Haberman B, and Verno A., "The new educational imperative: Improving high school computer science education," *Final Report of the CSTA Curriculum Improvement Task Force*, Feb 2005.
- [2] Sotirofski K, Kukeli A, and Kalemi E., "Challenges Of Teaching Computer Science In Transition Countries: Albanian University Case," *Journal of College Teaching and Learning*, vol. 7, no. 3 pp. 79, March 2010.
- [3] Rebecca A. Clay, "What you should know about online education," *American Psychological Association*, vol. 43, no. 6, pp. 42, June 2012. (accessed Jun 30, 2016).
<http://www.apa.org/monitor/2012/06/online-education.aspx>.
- [4] JISC, "E-Learning Pedagogy Programme," *E-learning pedagogy JISC*, <https://www.jisc.ac.uk/rd/projects/e-learning-pedagogy>. (accessed June 20, 2016).
- [5] R. Sims, "Rethinking (e) learning: A manifesto for connected generations," *Distance Education*, vol. 39, no. 92, pp. 153-164, August 2008.
- [6] KPESE, "IT Projects, Elementary & Secondary Schools," *KPESE*, <http://www.kpese.gov.pk/>. (accessed June 30, 2016).
- [7] L. G. Muradkhanli, "Blended learning: The integration of traditional learning and eLearning," *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*, Baku Azerbaijan, pp. 1-4, 2011.
- [8] Anderberg E, Nordén B, and Hansson B, "Global learning for sustainable development in higher education: recent trends and a critique," *International Journal of Sustainability in Higher Education*, 18; vol. 10, no. 4, pp. 368-78, Sep. 2009.

-
- [9] Drake SM and Burns RC. Meeting standards through integrated curriculum. *ASCD*, 2004.
- [10] Herazo Rivera JD, "Using a genre-based approach to promote oral communication in the Colombian English classroom," *Colombian Applied Linguistics Journal*, vol. 14, no. 2, pp. 109-26, Dec., 2012.
- [11] Fogarty R, "Ten ways to integrate curriculum," *Educational leadership*, vol. 49, no. 2, pp. 61-5, Oct. 1991.
- [12] Cooper S, Dann W, Pausch R. Teaching objects-first in introductory computer science. *In ACM SIGCSE Bulletin 2003 Feb 19* (Vol. 35, No. 1, pp. 191-195), ACM.
- [13] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
- [14] Reem Alebaikan and Salah Troudi, "Blended Learning in Saudi Universities: Challenges and Perspectives," *ALT-J, Research in Learning Technology*, vol. 18, no. 1, pp. 49-59, March 2010.
- [15] <https://www.getapp.com/education-childcare-software/education-elearning/>, accessed on January 2018.
- [16] <http://blog.ed.ted.com/2015/09/19/25-awesome-apps-for-teachers-recommended-by-teachers/>, Accessed on January 2018.
- [17] Hisam Barakat Hussein, "Attitudes of Saudi Universities Faculty Members towards using Learning Management System (Jusur)," *The Turkish Online Journal of Educational Technology*, vol. 10, no. 1, pp 43-53. April 2011.
- [18] <https://www.moe.gov.sa/en/TheMinistry/Education/Pages/InitiativesandProjectsofTheMinistryofEducation.aspx> AlMegren, A., & Yassin, S. Z. (2013).

Holy Qur'an Speech Recognition System Distinguishing the Type of prolongation

Bilal Yousfi¹, Akram M. Zeki², Aminah Haji¹

Abstract:

The act of learning and teaching of the Holy Quran has become a scientific practice to Muslims around. The stakeholders are faced with a huge challenge when it comes to the principle of application of Tajweed (that is, the rules guiding the pronunciation during the recitation of the Quran). There are several efforts made by previous systems on the development of feasible guiding techniques to the act of Tajweed. Unfortunately, liking the major control variables of the practices of Tajweed in those approaches were neglected. In order to fill this gap, this research presents a speech recognition system that distinguishes the types of Madd (elongated tone) or prolongation and the type of Qira'at (method of recitation) related to Madd. The proposed system is capable of recognising, identifying, pointing out the mismatch and discrimination between two types of Madd namely, The greater connective prolongation and The Exchange Prolongation rules for Hafss and Warsh for the verses that contains the two rules, that were made by the expert found in a database. Furthermore, this study used Mel-Frequency Cepstral Coefficient (MFCC) and Hidden Markov Models (HMM) as feature extraction and feature classification respectively.

Keywords: *Holy Quran; Tajweed; Qira'at; Sound, Mel-Frequency; Hidden Markov Models.*

1. Introduction

The Holy Qur'an was revealed with Tajweed rules, and it's important for readers to apply those rules during recitation. According to Qira'at science, each Qira'at has its own rules of Tajweed. Table 1 shows the difference between Hafss and Warsh in terms of the greater connective prolongation and the exchange prolongation. The act of Tajweed is the body of knowledge perfecting and laying the path to the understanding of the articulation of the Holy Quran letters and reaching the utmost level in pronouncing them properly. It is quite obvious that the application of the rules of recitation of Holy Quran can be performed by giving every letter of the Qur'an its rights and dues. Various lexical characteristics emerge when reciting the Qur'an and observing the rules that apply to those letters in different situations. For instance, an onwards break and the need for reciters to halt for either

a short period of time or longer. This is a unique attribute that influences people. This tends to be habitual and in some certain situation, even without due consideration of rules, it becomes an important part of learning the Holy Quran. However, the facts still lie with variation of short or long period needed. Many other features and rules that need to be taken in to consideration make it necessary for researchers to formulate techniques to ease the learning and teaching the act of Tajweed.

The act of Tajweed brings along some well-defined rules of recitation of Al-Quran. Noticeably, those rules create a big difference between normal Arabic speeches and the Quranic verses. Madd or prolongation is one of the significant tajwid rules. It stands for extending or prolonging sound with a letter of the Madd. It is divided into two groups:

The Original Madd and The Secondary Madd. The later presentations do not involve a

¹ Kulliyah of Information and Communication Technology (KICT)

² International Islamic University Malaysia (IIUM)

Corresponding Email: : yousfi.bilal@hotmail.fr

hamzah before it, and there should not be a hamzah or sukoon after it. Whereas the former has a longer timing (or the possibility of longer timing) than that of the natural Madd. This comes into being because of the present of a hamzah or a sukoon before and after it. Considering these presentations and with the availabilities of two rules of prolongation (mudud);the greater connective prolongation and the exchange prolongation rules under The Secondary Madd. This research focuses on their dynamics.

The Exchange Prolongation: rule substitute prolongation which occurs when a hamza (ء) precedes a half Madd (ا or ي or و). This Madd is only found within one word and occurs when the hamza has the respective diacritic on it, for example. if the harf Madd 'waaw' follows a hamza, the hamza has a dammah on it [1].

The Greater Connective Prolongation : rule of the greater connective prolongation occurs If the pronoun/possessive pronoun (هـ) is at the end of a word and it has a vowel of a dhammah or a kasrah, is between two voweled letters, and the first letter of the next word is a hamzah, it is permissible to lengthen according to the type of recitation [1].

The rules of prolongation are directly attached to Tajweed rules, these are mostly studied independently, however, their correct application is most reliable when performing subjective assessment in the present of Tajweed experts. That is, where experts of Tajweed are involved in the treatment of the rules governing the Tajweed application. Unfortunately, it's difficult to get experts at any time while in Quranic learning and teaching application practices. In most cases, many people will like to find Tajweed expert who will listen to them and point out mistakes if any while in Quranic learning session. Thus, it has become an important task to develop a learning software that will aid the practices of Tajweed in the act of learning Quran. This will guide people to practise reading of the Quran in correct way of Mudud spelling. Crucial to this is utilizing a speech recognition system for

distinguishing the type of Madd as well as the type of Qira'at.

The remaining part of this paper is organized as follows. The current section is section one, followed by section two which represents this research related work. Section three describes the speech recognition approach, section four present the research methodology and section five presents the outcomes of this research. Finally, section six presents the conclusion.

TABLE 1. The difference between the Greater Connective Prolongation and the Exchange Prolongation according to Hafss and Warsh Qira'at.

The Type of mudd	Hafss	Warsh
The Exchange lengthening مد البدل	lengthened 2 counts	lengthened 2, 4, or 6 counts
The greater connective prolongation م د الصلة الكبرى	lengthened 4 or 5 counts	lengthened 6counts

2. Previous Work

Previously, great effort has been made mostly for the study of Holly Quran Arabic speech recognition. There are many reviews of the evolution of Arabic Holy Qur'an ASR. Several theories were proposed for evaluating high accuracy for Arabic Qur'an speech recognition [2].

One of the most important researches in this area presented by [3]. The use of Multilayer Perceptron for classification of the pronunciation of Qalqalah Kubra (a Tajweed rule) has been presented. Feature extraction technique using MFCC has been utilized to extract the characteristics from Quranic verses' recitation. The technique used was able to achieve recognition rate within the range of (95% to 100%). Thus, the study has contributed to identifying correct and incorrect Qalqalah Kubra pronunciation.

The needs for people to be checking Tajweed rules in Quran verses by using

interactive way of learning without the guidance and presence of an expert has been emphasized in [4]. Thus, the design, implementation and evaluation of an automated Tajweed checking rules engine for Qur'anic learning were also presented. This has been validated with MFCC features and HMM model, where an accuracy of 91.95% (ayates) and 86.41% (phonemes) were obtained.

Generally, people were mostly involved in independent practice of listening of recitation of famous reciters with the aim of learning from that. This is appropriate but Makhraj might be missing, thus a novel technique based on correct Makhraj has been proposed in [5]. The technique was evaluated with MFCC as feature extraction and Mean Square Error (MSE) as a pattern matching technique. Thus, accuracy of the approach based on False Reject Rate (FRR) and Wrong Recognition (WR) has been obtained, where the percentage of FRR for all recitation is 0% and the accuracy of the system is 100%.

A novel model which distinguishes the type of recitation of holy Quran has been proposed in [6]-[7]. Feature extraction technique used is Mel-frequency Cepstral Coefficients (MFCC). Hidden Markov Model (HMM) is also employed for classification. Similar to other techniques, this has also aimed at improving learners techniques of recitation of the Holy Quran [8]-[9].

3. Speech Recognition for Distinguishing the type of mud according HAFSS or WARSH

Speech recognition technique involve the process of recording speech or acoustic signal which will be accurately and efficiently convert into a set of words [10]. The steps involve producing a speech recognition system is presented in Figure 1.

The aim of ASR is to extract, characterise, and recognise, the information about speech identification. The system consists of three basic stages as shown in fig 1.: pre-processing, where the recording speech (verses) signals is

passed through the pre-processing block to remove the noise and separate desirable voice from undesirable once and detect the start point and end points of verses.

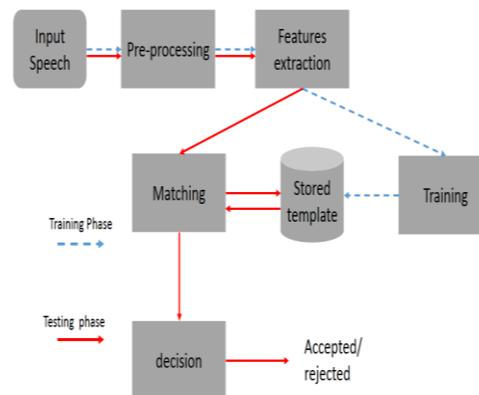


Fig. 1. Block Diagram of speech Recognition System.

Feature extraction is the process of extracting parameters that are unique to each word from the input sample of speech. This can be used to differentiate between a wide set of distinct words. The Mel-Frequency Cepstral Coefficient (MFCC) is considered the most evident example of a feature set [11]. This is widely utilized in speech related studies. However, it is closely related to the logarithmic perceptual ability of the humans. As a way to extract the coefficients, the speech sample is taken as the input. Pre-emphasis is applied to pass the signals through a filter which emphasises higher frequencies. This process will increase the energy of signal at higher frequency. After pre-emphasis, the signals are directed for frame blocking and windowing. Frame blocking is the process of segmenting the speech samples obtained into frames with the length within the range of 20 to 40 msec of N samples, with adjacent frames. Windowing is aimed at minimising the discontinuities of a signal at the beginning and at the end of each frame. This step follows by converting each frame from the time domain into the frequency domain by utilizing DFT. Then to generate the Mel filter bank. This is

done by a set of triangular filters that are used for each frame with actual frequency with Mel frequency as middle frequency. The triangular filter represents the process of Mel scaling in the signal. The next step is the computation of logarithmic of signal energy. The goal of logarithmic signal energy process is to adapt with the system just like human ear. In order to obtain the MFCC, the result of energy logarithmic is processed with Discrete Cosine Transform (DCT). Equation 1 present the approximate empirical relationship to compute the Mel frequencies for a given frequency f expressed in Hz:

$$\text{Mel}(f) = 2595 * \log_{10}(1 + f/700). \quad (1)$$

Figure 2 shows the steps involved in MFCC feature extraction.

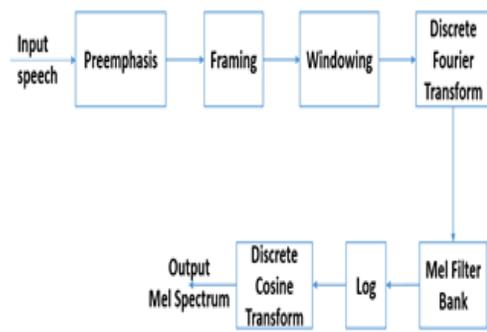


Fig. 2. Block diagram of the computation steps of MFCC.

The features classification or pattern recognition is the process of identifying similarities of spoken words between an extract feature from the input signal and set of acoustic models stored in the database. HMM [12] techniques were used by many researchers for speech recognition.

4. Methodology

The research methodological approach focus on prolongation type recognition and is presented in Figure 3.

The first step involves data collection of the Quranic recitation samples from different experts Qari (Reciter). These experts were known to have Ejazah in Hafss and Warsh.

Each of them recite specific verses that contained the two kind of prolongation (mad): The greater connective prolongation and The Exchange Prolongation rules for Hafss and Warsh. The samples were for collected many times in correct way. Thus, the entire samples are stored as the raw dataset that are prepared for pre-processing.

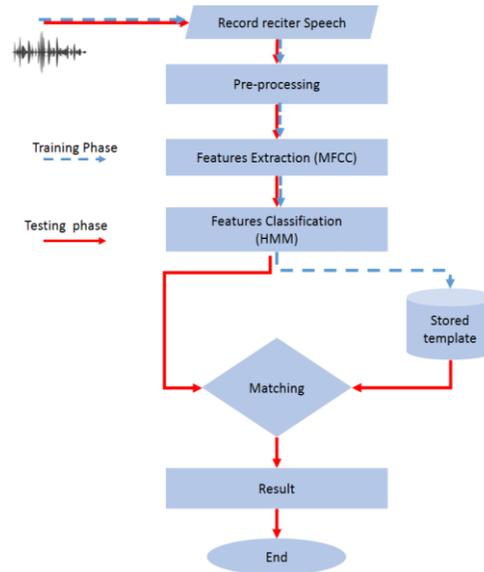


Fig. 3. Prolongation type identification for Qur'an flow chart.

The collected raw data are pre-processing to remove the noise contained in the speech signal and separate the desirable voice from undesirable once. This reduce the group of attributes which assure only the information that wants to be conveyed. The MFCC algorithm is applied to extract and generate features vector which is extensively used as an input for recognition purposes. The classification is done by HMM model. It calculates the HMM parameters. Training phase is characterised by extracting features using large number of samples "training data", and testing phase is characterised by extracting features from testing data "data speech". Testing data (the user recorded) are matched with voice features stored in the database, to

provide responses based on whether they recited correctly or incorrectly. Then the comparison acknowledges the users' level of accuracy. A codebook models (stored template) in the database that is constructed from training data is used for the experimental records.

5. Experiment and Results

An experiment was carried out intended to present some of the recognition scenarios. The acoustic model of some Holy Qur'an verses speech signal contained the two kinds of prolongation were used to show the differences between that exist among them. This is based on the different type of recitation of Hafss and Warsh as shown in Figure 4, 5 and 6, and Figure. 7 and Figure. 8. The recognition was carried out based on the guidelines presented in section 4.

This research focus on five words from verses of the holy Qur'an. These verses have been chosen for each of the greater connective prolongation and The Exchange Prolongation as shown in table 2 and table 3, which was recited by the two famous types of Qira'at; namely the Qira'at of Hafss from Asim and the Qira'at of Warsh from Nafi.

TABLE 2. Verses selected for the Greater Connective Prolongation.

VERSES	Surah
"خَيْرَانَ لَهُ أَصْحَابٌ [6:71]"	<i>Al-An'am</i> الإنعام
"يَدْعُونَهُ إِلَى الْهُدَى [6:71]"	<i>Al-An'am</i> الإنعام
يُشْرِكْ بِعِبَادَةِ رَبِّهِ أَحَدًا [18:110]	<i>Al-Kahf</i> الكهف
أَيُّحْسِنُ أَنْ لَمْ يَزِدْهُ أَحَدٌ [90:7]	<i>Al-Balad</i> البلد
فَتَمَّ مِيقَاتُ رَبِّهِ أَرْبَعِينَ [7:142]	<i>Al-Araf</i> الأعراف

The experimental test and results of this research is presented in this section through the following:

- Data collection or training phase.
- Recognition phase.

The first phase involves collection of the recitation of the sample data, that will aid in

extracting and training the features. The data are found at the Reciter's database. The database is selected from Internet. Verses of the Holy Qur'an for each of exchange prolongation and the greater connective prolongation are the key objects used. This are tested on the most popular reciters such as Sheikh Al-Hosry. Five (5) verses are recited by three (3) reciters with two (2) categories of Qira'at which are Warsh and Hafss on The Exchange Prolongation as well as The Greater Connective Prolongation respectively. A total of sixty (60) of data samples are obtained. All the samples are passed through the extraction stage in order to extract and represent the features in the form of frequency on Mel scale. Delta coefficients of Mel Coefficients are calculated and then, trained and recognized using HMM. This are used as reference patterns and stored as Reciter's database.

Recognition phase involves verifying the recitation of new reciters to the pre-stored value against the entire reciter in reciter's database. The results are tested against the specified objectives of proposed system. The developed system is tested by performing the MFCC algorithm for features extraction from the Qur'anic recitation of samples data used and then, matching/testing against the trained HMM model of data templates, using the same classification of HMM method. The HMM algorithm is anticipated to get the best results of identification system.

The recognition accuracy rate is calculated using equation 2:

$$\text{Accuracy} = (\text{number of correct samples} / \text{total samples}) \times 100 \quad (2)$$

The experimental results of the testing process are presented here. The experiment reveals the extracted features of 10 verses of the Qur'anic recitation which were directly compared with the data based on the Model. As a result, the test result on the training data obtained for this study is at 60% and 50% for The Exchange Prolongation according to Hafss and Warsh and 40%, 70% for the greater connective prolongation according to Hafss and Warsh respectively (see Table 4).

The gathered results have shown some enhancements compared to the previous findings. The research contributions lie with the improve performance and efficiency of the proposed technique.

Although, the system faces some drawbacks, with the extra noise due to audio file compression and poor quality during the recording process. Yet, high-performance measure was achieved.

TABLE 3. The verses selected for the Exchange Prolongation.

VERSES	Surah
“يَا أَيُّهَا الَّذِينَ آمَنُوا اذْكُرُوا اللَّهَ ذِكْرًا كَثِيرًا [33:41]”	<i>Al-Ahzab</i> الاحزاب
“ وَمَا تَشَاؤُونَ إِلَّا أَنْ يَشَاءَ اللَّهُ رَبُّ الْعَالَمِينَ [81:29]”	<i>Al-Takwir</i> التكوير
“الْمُؤْمِنِينَ لِيَزِدُوا إِيمَانًا مَعَ إِيْمَانِهِمْ” [48:4]	<i>Al-Fath</i> الفتح
“وَإِنَّ الَّذِينَ أُوْتُوا الْكِتَابَ لَيَعْلَمُونَ أَنَّهُ الْحَقُّ مِنْ رَبِّهِمْ [2:144]”	<i>At-Tawba</i> التوبة
“وَجَاؤُوا أَبَاهُمْ عِشَاءً يَبْكُونَ [12:16]”	<i>Yusuf</i> يوسف

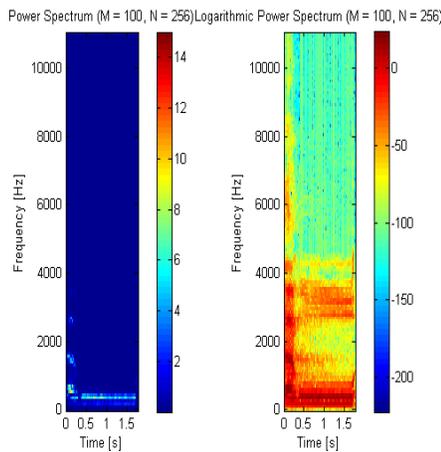


Fig. 4. Power spectrum plot of spoken word له خَيْرَانِ لَهُ أَصْحَابٌ the Greater Connective Prolongation According to Warsh.

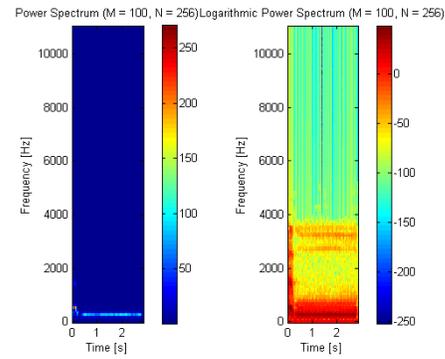


Fig. 5. Power spectrum plot of spoken word له خَيْرَانِ لَهُ أَصْحَابٌ The Greater Connective Prolongation According to Hafss.

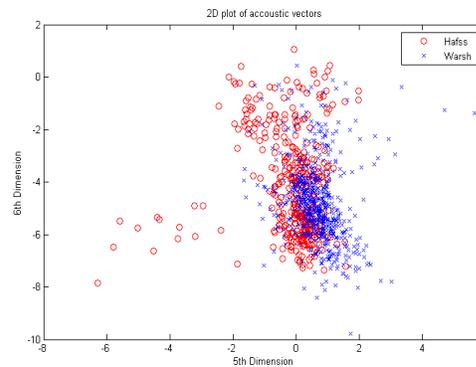


Fig. 6. 2D Plot of acoustic vector of spoken word له خَيْرَانِ لَهُ أَصْحَابٌ The Greater Connective Prolongation According to Warsh and Hafss.

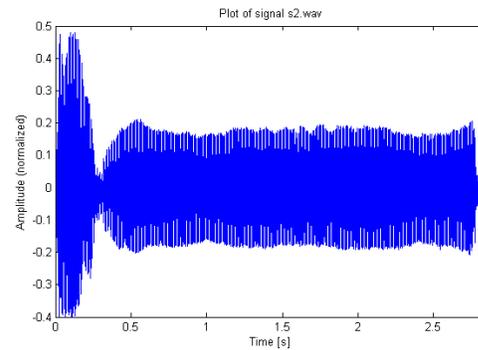


Fig. 7. Speech signals of spoken word له خَيْرَانِ لَهُ أَصْحَابٌ The Greater Connective Prolongation According to Warsh.

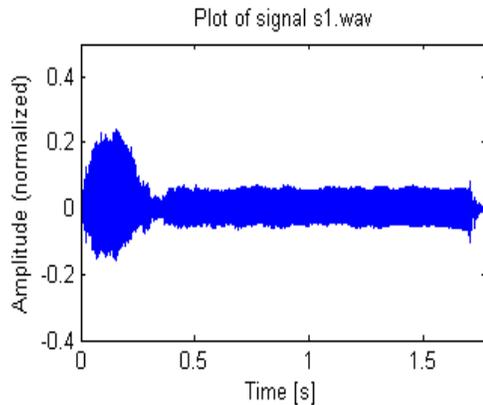


Fig. 8. Speech signals of spoken word خَيْرَانَ لَهُ، له "أصْحَابُ" The Greater Connective Prolongation According to Hafss.

TABLE 4. Model tuning results.

Prolongation type	The Exchange Prolongation		The Greater Connective Prolongation	
	Warsh	Hafss	Warsh	Hafss
<i>Qira'at type</i>	Warsh	Hafss	Warsh	Hafss
<i># of utterances</i>	10	10	10	10
<i>Correct</i>	06	05	04	07
<i>Wrong</i>	04	05	06	03
<i>% Accuracy</i>	60%	50%	40%	70%

Figure 9 shows the recognition accuracy rate of each kind of prolongation type where y-axis contains results and x-axis contains the types of Madd.

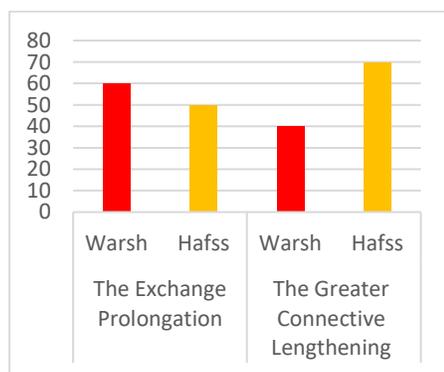


Fig. 9. The accuracy rate of proposed system.

6. Conclusion

This paper has developed on theory and practice based for developing a high-performance Tajweed system that assist in proper Tajweed Qur'anic recitation based on the automatic speech recognition system. The research utilized, Mel-frequency Cepstral Coefficients (MFCC) and HMM (Hidden Markov Model) algorithms to enable the validation of the proposed system. Several experiments were carried out. The experimental results on a database indicate that the feature extraction method and recognition method used for this research appropriate for Arabic recognition system are feasible.

There are other several techniques such as Liner Predictive Coding (LPC) and Artificial Neural Network (ANN) could also be used for similar research approach. The findings from those might be different, therefore, this research recommend future work to focus on using discriminative training techniques which might improve the discrimination between some confusable pronunciation alternatives.

ACKNOWLEDGMENT

The authors would like to thank the Research Management Centre and the Faculty of Information and Communication Technology, the International Islamic University Malaysia for their supports.

REFERENCES

- [1] K. C. Czerepinski and A. D. A. R. Swayd, Tajweed Rules of the Qur'an. *Dar Al-Khair Islamic Books Publisher*, 2006.
- [2] B. Yousfi and A. M. Zeki, "Automatic Speech Recognition for the Holy Qur'an, A Review," in *The International Conference on Data Mining, Multimedia, Image Processing and their Applications (ICDMMIPA2016)*, 2016, p. 23.
- [3] H. A. Hassan, N. H. Nasrudin, M. N. M. Khalid, A. Zabidi, and A. I. Yassin, "Pattern classification in recognizing Qalqalah Kubra pronunciation using

- multilayer perceptrons,” *IEEE symposium on Computer Applications and Industrial Electronics (ISCAIE)*, 2012, pp. 209–212.
- [4] D. Raja-Jamilah Raja-Yusof , Fadila Grine, N. Jamaliah Ibrahim, M. Yamani Idna Idris, Z. Razak, and N. Naemah Abdul Rahman, “Automated tajweed checking rules engine for Quranic learning,” *Multicult. Educ. Technol. J.*, vol. 7, no. 4, pp. 275–287, 2013.
- [5] A. N. Wahidah et al., “Makhraj recognition using speech processing,” *7th International Conference on Computing and Convergence Technology (ICCCT)*, 2012, pp. 689–693.
- [6] B. Yousfi and A. M. Zeki, “Holy Qur’an speech recognition system distinguishing the type of recitation,” *7th International Conference on Computer Science and Information Technology (CSIT)*, 2016, pp. 1–6.
- [7] B. Yousfi, A. M. Zeki, and A. Haji, “Holy Qur’an Speech Recognition System Mudud Tajweed Rule Checking,” *Int. J. Islam. Appl. Comput. Sci. Technol*, pp. 10–18, 2016.
- [8] B. Yousfi and A. M. Zeki, “Holy Qur’an speech recognition system Imaalah checking rule for warsh recitation,” *IEEE 13th International Colloquium on Signal Processing & its Applications (CSPA)*, 2017, pp. 258–263.
- [9] B. Yousfi, A. M. Zeki, and A. Haji, “Isolated Iqlab checking rules based on speech recognition system,” *8th International Conference on Information Technology (ICIT)*, 2017, pp. 619–624.
- [10] N. Zerari, B. Yousfi, and S. Abdelhamid, “Automatic Speech Recognition: A Review,” *Int. Acad. Res. J. Bus. Technol.*, vol. 2, no. 2, pp. 63–68, 2016.
- [11] S. B. Davis and P. Mermelstein, “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences,” in *Readings in speech recognition*, Elsevier, 1990, pp. 65–74.
- [12] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

Analysis on Energy Efficient Protocols-Wireless Sensor Networks

Iqra Tariq¹, Talal Bin Maqsood¹, Babur Hayat Malik², Mareena Asghar²,
Quratulain Gulzar²

Abstract:

Wireless Sensor Networks are among the networks which have the aptitude to be used in harsh surroundings. Wireless Sensor networks utilize micro sensor nodes. Although, sensors offer high quality and proficiency to put up with faults but their inadequate battery life is indulging it in impediments. Because limited battery life hinders communication among nodes in Network. Keeping in view above particulars; analysis has been completed on protocols that resolve crisis of low energy due to limited battery. This paper states analysis on Low Energy Adaptive Clustering Hierarchy protocol that has now decidedly shaped into Advanced Low Energy Adaptive Clustering Hierarchy protocol to put off energy dissipation in improved manner. This Review paper is intended for comparing Energy Trends in Low Energy Adaptive Clustering Hierarchy, Advanced Low Energy Adaptive Clustering Hierarchy, and Low Energy Adaptive Clustering Hierarchy –Clustering. It is static and has heterogeneous routing protocol, Multi Hop and Distributed Energy Efficient Clustering Protocol. How Leach protocols could be improved. Result has offered availability of low power sensors consisting of sensor nodes that use clustering practice. Outcome of this analysis illustrates that utilization of energy can be minimized in protocols by facilitating equal load allocation among all nodes.

Keywords: LEACH protocol; Wireless Sensor Network (WSN); Multi Hop (MHT); Base Station (BS); Cluster Head (CH); Distributed Energy Efficient Clustering (DEE); Medium Access Control (MAC).

1. Introduction

The Wireless Sensor Network (referred to as WSN from now) has become an advance trend of today’s Networking era.

It is being used for past years for observing physical world communication. It is a particular type of ad hoc network that monitors substantial world by using undersized sensors. These sensors are hundreds or thousands in amount and are compactly or sparingly distributed in the network [1]-[2].

At the moment of Communication; application comprising sensor nodes have intentions to send gathered information by sensing target region [3]. After sensing target region, regular nodes launch their information

to desired CH which is accountable for transferring information to the BS. The distinctive configuration of wireless protocol is illustrated in Fig.1 below.

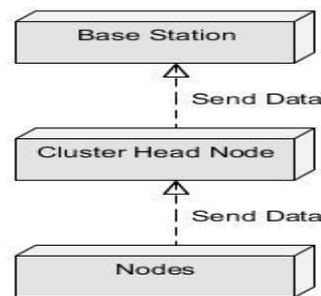


Fig. 1. Wireless Protocol Configuration.

¹ University Of Management and Technology, Sialkot Pakistan

² University of Lahore, Chenab campus, Gujrat Pakistan

Corresponding Email: iqra.tariq@skt.umt.edu.pk

Battery is set with a sensor which supplies power to the nodes in the network. This battery fails because it couldn't offer countless power in network [4]-[5]. For this purpose, Low Energy Adaptive Clustering hierarchy (LEACH) [6] Protocols are designed which consume smaller amount of energy and consequently boost network life span. This LEACH has been further enhanced and affirmed earlier. This paper enlightens progressions that have been made in field of Wireless Sensor Network which put off energy loss and add to network lifespan and provides comparative study among Advanced LEACH (Ad-LEACH) [6] protocol, MH-LEACH (Multi-Hop LEACH) [7]-[8] protocol and DEEC (Distributed Energy Efficient Clustering) [9] protocol correspondingly.

This paper is alienated as follows: Section 2 portrays Literature Survey, Section 3 portrays Latest Energy Efficient Protocols, and Section 4 presents Comparison between LEACH, LEACH-C, Ad-LEACH and Multi Hop and DEEC. Section 5 portrays conclusion which wrap up this Paper.

2. Literature Survey Of Traditional Energy Efficient Protocols

WSN has made today's communication cost effective, simple and easy. WSN supports heterogeneous applications [10]. But the problem that made WSN unreliable is short lifetime of its battery [11, 12]. There are many protocols that had been implemented in WSN but these traditional protocols were not able to optimize

- 1) Direct Communication
- 2) Minimum Transmission Energy
- 3) Multi-hop Routing
- 4) Static Clustering.

Every Protocol intended to resolve setbacks for WSN have been conferred underneath one by one.

A. LEACH and its limitations

Some students of Massachusetts Institute of Technology (MIT) took into account all these above mentioned problems and introduced a new protocol, called Low Energy Adaptive Clustering Hierarchy (LEACH) [13].

LEACH is presented in 2000 [13]. LEACH protocol consists of subsequent characteristics conversed below.

- 1) It is clustering based protocol. Clusters consist of nodes; these nodes acquire data from each node and hence fusing data to the cluster head by sending meaningful set of information from the nodes. Each CH broadcasts data to BS. Configuration of LEACH is the same as Simple Wireless Structure shown in fig. 1 [13-16].
- 2) It allocates Energy to the sensors of cluster head. LEACH is able to distribute energy dissipation evenly throughout the sensors that double the useful system lifetime, for the network [13-17].

A leach is able to solve most of the problems faced by traditional protocols. But there were flaws in LEACH itself that are`

- 1) Nodes can only send data to their cluster head which makes more likelihood for CH's to die quickly. Hence life-span of Network is minimized.
- 2) Nodes cannot communicate among each other. If a sensor is not a cluster head, afterward it cannot launch data to further sensors. Consequently, distance among nodes turn out to be larger which in outcome shortens energy level of entire Network

Above particulars make it obvious that network protocols which extend battery life are more purposeful. So, these flaws were eliminated by introducing two more advanced protocols Multi-Hop Low Energy Adaptive Clustering Hierarchy (MH-LEACH) and Advanced LEACH which are conversed in Section 3.

B. LEACH-C

LEACH-C is presented in 2002. LEACH-C (Low Energy Adaptive Clustering Hierarchy Centralized) uses clustering Centralized Technique for assembling clusters as discussed. Clusters assembly in LEACH-C

has CHs whose initiative is to execute subsequent roles [18-23].

- 1) CH's are there to determine locality and energy intensity of each node that is accountable for conveying information.
- 2) CHs send Meaningful Data to Base Station. LEACH-C executes tasks in Rounds. In every Round, BS ensures by estimating the average energy that energy is circulated uniformly amongst each node in WSN network [23].

In Centralized technique, Central unit has all the data collected from different nodes. Central network performs its responsibility for estimating each nodes position in network; whereas, in distributed technique, there are estimations of each node instead of having central unit and as a result, location is estimated grounded on local data collected from its adjacent nodes. The likelihood of bottleneck increases in centralized algorithm as connection of node can be lost if an error occurs or a critical node expires, whereas in distributed technique failure of one node doesn't affect system decisively [24]. Consequently distributed technique is more vigorous than Centralized Technique.

C. Ad-LEACH

Ad-LEACH (Advanced Leach Energy Efficient protocol) is presented in 2008 [25]. It has Static Clustering approach. BS position and arrangement of clusters of whole Network is predefined. After deciding Cluster Head, TDMA serves in data transmission as discussed below.

1) Cluster arrangement

Low Energy Adaptive clustering hierarchy launch network in form of static and unending clusters. In Ad-LEACH Square and Rectangular could be the shape of clusters according to the necessity and region offered. The review we carried illustrates both shape clusters.

Adjacent clusters would have segregated protocols in each cluster. For lessen complication and power dissipation, motivation is in separating entire region into

little static fields. Organizing immense field of operation is more difficult than clusters covering small portions. Therefore, Outcome of clusters formation lowers power intensity of their messages being transmitted [25]-[29].

2) Cluster Head Decision

After Cluster formation, it's time to decide cluster head because each cluster has separate Ad-LEACH protocols. CHs are selected on the basis of left over energy they contain while being in clusters. Nodes in network need prerequisite awareness of its entire energy and life-time of network in WSN. Networks entire power is transmitted from BS to each node in DEEC.

$$T(n) = \begin{cases} \frac{P_d}{1 - P_d * \left((r) \bmod \frac{1}{P_d} \right)} & \text{if } n \in A \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

The CH is chosen at every single round with the assistance of formula (1). The Threshold T (n) is generated by putting Percentage as 5% and A is nodes other than chosen CH's [5, eq (1)]. In the manner CH is elected, the elected cluster head must inform each node regarding its existence in clusters. CH and remaining nodes use Carrier Sense Multiple Access (CSMA), which is the protocol of MAC in the network.

3) Client Scheduling

Client information is received by CH from each node in cluster by building TDMA (Time division Multiple Access) schedule. Schedule is formed for all of its nodes which are used for data distribution towards Cluster Head (CH) Node.

4) Data Transmission

After building TDMA, broadcasting of data can take place. CH can be approached by its client nodes in only allocated time period. For the duration of Unallocated Time Period, client nodes are required to turn their radio off

for avoiding energy dissipation. Broadcasting energy level is based on (Received Signal Strength) RSS which is principally selected by every node itself. As there is load on CH, there is a possibility that CH might fail. CH Rotating is suggested in Ad-LEACH so that CH lifespan can't be endangered [29].

The difference between LEACH and ALEACH (Advanced low Energy Adaptive clustering hierarchy) is that in ALEACH network carry out broadcasting messages in rounds and elects CHs rather than identifying environmental position of each node. Then Multi-Hop Routing, DEEC is proposed in 2011 [30]-[31], 2014 [32], which is an expansion of above LEACH protocols which is discussed in this paper in section 3.

3. Latest Energy Efficient Protocols

A. Multi-Hop

In [4], Multi-Hop Communication is offered which is a new method of clustering. MH-LEACH have cluster Heads whose essential reason is to keep energy and which as outcome add to network life span. Energy dissipation is one of the most important aspects of MH-LEACH.

The algorithm implemented in MH-LEACH that transfers data to farthest location using lowest possible energy because in this protocol, a node sends data to its nearest node resulting in lower consumption and that makes a network reliable. In this algorithm, a sensor sends data to the base station only when it has received the data completely. In this protocol, sensors remain dormant when data is not being sent or received. MH-LEACH decides CHs in similar technique as LEACH protocol [31]. There are two manners in which communication is performed in Multi-Hop.

1) *Intra Cluster Transmission*

Inter Cluster Transmission is the type of Transmission which transmits information by gathering data. Clusters have CHs which obtain data from nodes and aggregate information for transferring the concluded information to BS.

2) *Intra Cluster Transmission*

Inter Cluster Transmission is the type of Transmission which transmits information by gathering data from all member of nodes.

There are two phases of MH-LEACH:

Phase 1: In first phase, cluster-headers are defined as a part of LEACH protocol. Then they make an announcement and all the cluster headers construct their routing table taking in account the level of signal received.

Phase 2: When the cluster header sends data, according to the routing table previously constructed, the base station checks whether the cluster head is not clashing with other routes. After checking this condition, base station sends data to another node.

Multi Hop Transmission makes clusters in Intra Communication which approaches CHs by transmitting data after gathering it from other member's nodes. It follows the same mechanism as LEACH does by executing tasks in Rounds. It fundamentally decides pathway that have minimum hops among CHs and Base Station (BS) [31].

B. DEEC

In WSN, Distributed energy efficient clustering algorithm (DEEC) is presented. DEEC reflect on attributes which are Heterogeneous [32]. DEEC algorithm progress scalability and decrease amount of battery utilization. Consequently, the Algorithm of choosing CH is pursued of DEEC in Ad-LEACH [33]-[34].

4. Comparison Between Traditional And Latest Energy Efficient Protocols On The Basis Of Their Properties

A. Protocol Properties discussion

Comparison of network design for Protocols on the basis of properties is prepared in this Section by keeping in view analysis of LEACH and other Latest Energy Efficient protocols already discussed in previous sections. The properties of protocols which are stated in the Table 1 are described underneath one by one.

1) *Life span of Network*

Lifespan is the characteristic which tells the duration of nodes that are alive and dead in the network.

2) *Energy Usage*

Energy Usage is the characteristic of WSN which determines the usage of energy while sending data from CH to BS.

3) *Scalability*

Scalability determines the amount of data that can be sent without any failure of nodes.

4) *Transmission of packets*

Packet Transmission rate is analyzed of the network.

5) *Pathway Choice*

Nodes select whether the path would be Single Hop or Multi Hop. In Single Hop Routing, nodes send data to Cluster Heads. CH's sends data to Base Station directly. In Multi-Hop Routing, nodes send data to Cluster Heads. CH's send data to nearest CH's and thus shorten the distance towards BS.

6) *Amount of active nodes*

Nodes chosen as CH have greater probability to die soon. This property analyzes

the Active Nodes and Dead Nodes measurement.

7) *Classification*

Classification is made on the basis of Hierarchical clustering. In Hierarchical Clustering, there are many clusters in entire network. Clusters have Cluster Heads. CH's are there to determine locality and energy intensity of each node that is accountable for conveying information. CHs send Meaningful Data to Base Station.

8) *Position Awareness*

The location is found of network in some protocols so that data can be gathered.

9) *Mobility*

Mobility determines base station kind which varies from fixed Base Station to changing Base Station. Normal Nodes send data to chosen CH. Then chosen CH's send useful information to BS.

10) *Data Aggregation*

Data Aggregation utilizes aggregating algorithm. The algorithm used for this intention is Centralized Algorithm. It is the process in which information is divided into packets to send consequential information to the BS in the network.

B. Results

TABLE 1. Comparison between Diverse Routing Protocol on basis of their Properties.

S. No	Protocol Properties/Reference	Routing Protocols in Wireless Sensor Network				
		<i>Leach</i>	<i>Leach-C</i>	<i>Ad-Leach</i>	<i>Multi-Hop</i>	<i>DEEC</i>
1.	Lifespan of Network/[2]	Great	Greater than Leach	Greatest as CH rotates	Improved than Leach	Improved
2.	Energy Usage/[35]	High	Unit Energy Less than Leach	More than LEACH	Reliable	Reliable
3.	Scalability/[2]	Inadequate	Inadequate/Very short	More than Leach	Adequate	Adequate/Enhanced

4.	Transmission of Packets/[35]	fewer	Extra Data for every Unit Time	More than Leach	More than Leach	More than Leach
5.	Pathway Choice/[2]	Single Hop	Single Hop	Single Hop	Multi-Hop	Multi-Hop
6.	Amount of Active Nodes/[35]	Identical	Nodes Die Earlier	Less Nodes Die	less but exploit more time to fail	More time to fail
7.	Classification/[35]	Proactive/ Hierarchical Clustering	Hierarchical Clustering	Hierarchical Clustering	Hierarchical Clustering	Proactive or Hierarchical Clustering
8.	Position Awareness/[2]	No	Yes	Yes	Yes	No
9.	Mobility/[2]	Unchanging Base Station	Unchanging Base Station	Unchanging BS	Change BS	Unchanging Base Station
10.	Data Aggregation/[35]	Yes	Yes	Yes	Yes	Yes

5. Conclusion

In WSN, Great piece of Research is being carried out with the intention to emphasize ways for sensor to save energy so that network life can be extended. Wireless Sensor Networks (WSN) set up ad hoc networks. Ad hoc networks agree to observing physical world through assistance of small sensors, which are sparingly or heavily distributed. Thus, DEEC and Multi-Hop protocol has the ability to be used in different applications. These protocols usage can consequence in controlling Energy Dissipation and progressing Networks Life Span which can become the basis of efficient Transmission in WSN networks. Review completed in this paper offers improved performance along with eradication of energy dissipation.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science

and Technology (No. NRF-2015R1D1A1A01059484).

REFERENCES

- [1] P. Bansal, P. Kundu, and P. Kaur "Comparison of LEACH and PIGASIS Hierarchical Routing Protocols in Wireless Sensor Networks," *International Journal of Recent Trends in Engineering and Technology*, vol. 11, pp. 139-144, June 2014.
- [2] J. Shen, A. Wang, C. Wang, Y. Ren, and J. Wang, "Performance Comparison of Typical and Improved LEACH protocols in Wireless Sensor Network," *1st International Conference on Computational Intelligence Theory, Systems and Applications*, Dec. 2015, pp. 187-192.
- [3] A. Singh, S. Rathkanthiwar, and S. Kakde, "LEACH Based Energy Efficient Routing protocol for Wireless Sensor Networks," *IEEE International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2016, pp. 1.

- [4] H. Dhawan, S. Waraich, "A Comparative Study on LEACH Routing Protocol and its Variants in Wireless Sensor Networks: A Survey," *International Journal of Computer Applications*, vol. 95, no. 8, pp. 21-27, June 2014.
- [5] M. S. Ali, T. Dey, and R. Biswas, "Aleach: Advanced leach routing protocol for wireless micro sensor networks," *Proc. of IEEE International Conference on Electrical and Computer Engineering (ICECE)*, 2008, pp. 909–914.
- [6] V. P. Tank and T. V.Vyas, "Comparison and Performance Analysis of Wireless Sensor Network Protocols LEACH and LEACH- using NS-2 Tool," *IEEE proceeding, IJECT*, vol. 3, no. 3, pp. 254-258, 2012.
- [7] A. Yektaparast, F.H.Nabavi, and A.Sarmast, "An Improvement on LEACH protocol (Cell-LEACH)," in *14th International Conference on Advanced Communication Technology (ICACT)*, Feb. 2012, pp. 992-996.
- [8] S. Faisal, N. Javaid, A. Javaid, M. A. Khan, S. H. Bouk, and Z. A. Khan, "Z-SEP: Zonal-Stable Election Protocol for Wireless Sensor Networks," *J. Basic Appl. Sci. Res.*, vol. 3, no. 5, pp. 132-139, 2013.
- [9] M. Y. Khan, N. Javaid, M. A. Khan, A. Javaid, Z. A. Khan, and U. Qasim, "Hybrid DEEC: Towards Efficient Energy Utilization in Wireless Sensor Networks," *World Applied Sciences Journal*, vol. 22, no. 1, pp. 126-132, 2013.
- [10] L. Qing, Q. Zhu and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer Communications*, pp 2230-2237, 2006.
- [11] Tahir, N. Javaid, A. Iqbal, Z. A. Khan, and N. Alrajeh, "On Adaptive Energy Efficient Transmission in WSNs," *International Journal of Distributed Sensor Networks*, vol. 2013(2013), Article ID: 923714.
- [12] M. B. Rasheed, N. Javaid, A. Javaid, M. A. Khan, S. H. Bouk, and Z. A. Khan, "Improving Network Efficiency by Removing Energy Holes in WSNs," *J. Basic Appl. Sci. Res.*, vol. 3, no. 5, pp. 253-261, 2013
- [13] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy Efficient Communication protocol for wireless micro sensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on. IEEE*, 2000, pp. 10.
- [14] R. ShanthaSelvaKumari, A. Chitra, and M. B. Devi, "Efficient -2 Level Energy Heterogeneity Clustering Protocols for Wireless Sensor Networks," *IEEE proceeding IJST*, vol. 9, no. 3, ISSN 0974-6846, February 2016, pp. 1-6.
- [15] K. Pawar, V. Pawar, "Enhancement of LEACH protocol Using Energy Heterogeneity Concept," *IEEE proceeding, IJETTCS*, vol. 2, no. 1, 2013, pp. 49-56.
- [16] Y. Mishra, A. Singhadia, and R. Pandey, "Energy Level Based Stable Election Protocol in Wireless Sensor Network," *IEEE proceeding IJETT*, vol. 17, no. 1, ISSN2231-5381, November 2014, pp.32-38.
- [17] M. Islam, Matin, and Mondol, "Extended Stable Election Protocol (SEP) for Three Level Hierarchical Heterogeneous WSN," *IEEE proceeding IET*, 2012, pp.1-4.
- [18] R. Pal, R. Shindhu, and A. K. Sharma, "SPE-E (RCH): Enhanced Stable Election Protocol Based on Redundant Cluster Head Selection for HWSNs," *IEEE proceeding SITE*, vol. 115, no. 5, 2013, pp. 104-114.
- [19] W. Li, F. C. Delicato, P. F. Pires, and Y. C. Lee, "Efficient Allocation of Resources in Multiple Heterogeneous Wireless Sensor Networks," *Elsevier, JPDC*, 2014, pp. 1775-177.
- [20] D. P. Nudurupati and R. K. Singh, "Enhancing Coverage Ratio using Mobility in Heterogeneous Wireless

- Sensor Networks,” *CIMTA, Elsevier*, 2013, pp. 538-545.
- [21] M. Akbar, N. Javaid, A. A. Khan, Z. A. Khan, and U. Qasim, “On Modeling Geometric Joint Sink Mobility with Delay-tolerant Cluster-less Wireless Sensor Networks,” *4th IEEE Technically Co-Sponsored International Conference on Smart Communications in Network Technologies (SaCoNet'13)*, 2013, Paris, France.
- [22] B. A. Attea and E. A. Khalil, “A New Evolutionary Based Routing Protocol for Clustered Heterogeneous Wireless Sensor Networks,” *Elsevier, ASC*, 2012, pp. 1950-1957.
- [23] P. Nayak and P. Shree “Comparison of Routing Protocols in WSN Using NetSim Simulator: LEACH vs. LEACH-C,” *IEEE proceeding, IJCA*, vol. 106, no. 11, ISSN 0975-8887, November 2014.
- [24] M. Aslam, N. Javaid, et al., “Survey of Extended LEACH-Based Clustering Routing Protocols for Wireless Sensor Networks,” *5th International Symposium on Advances of High Performance Computing and Networking (AHPCN-2012) in conjunction with 14th IEEE International Conference on High Performance Computing and Communications (HPCC-2012)*, June 2012.
- [25] Iqbal, M. Akbar, N. Javaid, S. H. Bouk, M. Ilahi, and R. D. Khan, “Advanced LEACH: A Static Clustering-based Heterogeneous Routing Protocol for WSNs,” *Pakistan*, 2013, pp. 1.
- [26] S. Basagni, M. Y. Naderi, C. Petrioli, and D. Spenza, “Wireless sensor networks with energy harvesting,” *Mobile Ad Hoc Networking: Cutting Edge Directions Inc*, 2013, pp. 703-736.
- [27] K. Huang, “Spatial throughput of mobile adhoc networks powered by energy harvesting,” *IEEE Transactions on Information Theory*, vol. 59, no.11, pp. 7597-7612, 2013.
- [28] G. Martinez, S. Li, and C. Zhou, “Wastage-aware routing in energy-harvesting wireless sensor networks,” *IEEE Sensors Journal*, vol. 14, no. 9, pp. 2967-2974, 2014.
- [29] G. Xiaoying and B. Chen, “A Novel Sensing Scheme for Dynamic Multichannel Access,” *IEEE Transaction on Vehicular Technology*, vol. 61, no. 1, pp.208-221, Jan. 2012.
- [30] Z. Su, Q. Xu, and Q. Qi, “Big Data in Mobile Social Networks: AQoE Oriented Framework,” *IEEE Network*, vol. 30, no. 1, pp. 52-57.
- [31] R. V. Biradar, D. Sawant, D. Mudholkar, and D. Patil, “Multi-hop routing in self-organizing wireless sensor networks,” *IJCSI International Journal of Computer Science*, vol. 8, no. 1, pp. 154–164, 2011.
- [32] J. Neto, A. Rego, A. Cardoso, and J. Celestino, “MH-LEACH: A Distributed Algorithm for Multi-Hop Communication in Wireless Sensor Networks,” in *ICN2014: The Thirteenth International Conference on Networks*, 2014, pp. 55-61.
- [33] J. H. Brand aoNeto, and Antoniel da Silva Rego, “MH-LEACH: A Distributed Algorithm for Multi-Hop Communication in Wireless Sensor Networks,” *The Thirteenth International Conference on Networks, Copyright (c) IARIA*, 2014. ISBN: 978-1-61208-318-6.
- [34] N. Javaid, T.N. Qureshi, A.H. Khan, and A. Iqbal “EDDEEC: Enhanced Developed Distributed Energy Efficient Clustering for Heterogeneous Wireless Sensor Networks,” *Elsevier, SEIT*, 2013, pp. 914-919.
- [35] M. Angurala and Bharti, “A Comparative Study between LEACH and PEGASIS- A Review,” 978-9-3805-421-2/16/\$31.00, IEEE, pp. 3272, 2016.

Analysis on Security Methods of Wireless Sensor Network (WSN)

Murtaza Ahmed Siddiqi¹, Abdul Aziz Mugheri², Mohammad Khoso²

Abstract:

Security has always been a major area of concern for WSN. Due to limited resources and size constraints of a node, WSN still lacks a comprehensive security mechanism for its operations. In this paper, some of the proposed security methods for WSN are being analyzed for the issues that still exist in the proposed security method. To perform the analysis on security methods some of the documented or implemented security algorithm by researchers are being studied and the issues with those algorithms are being highlighted. After performing the analysis, it is quite clear that most of the algorithm being proposed by researchers for WSN need to be designed keeping in view resources constraints of WSN.

Keywords: WSN; security; wireless; encryption.

1. Introduction

Wireless sensor network (WSN) is emerging as one of the most prominent and promising technology for numerous area. Its application areas including medical, industrial, agricultural, home appliance and military applications. WSN covers a broad domain of applications. That is why researchers are putting in a lot of efforts to achieve perfection in this technology. WSN can be explained as a network of (possibly very small with limited power and processing ability) devices identified as *nodes*. These nodes can be used to sense the environment (e.g temperature, air pressure) and can be used for multidimensional data gathering purposes. These nodes deliver the information gathered from field to the sink node using wireless links, this wireless communication can be multiple hops or directly relay to the sink node. Once the data is aggregated by sink node then as per requirement data can be relayed to the user using a gateway node, base station or at times direct access to a WSN node [1]. Since all this communication is carried out wirelessly, it brings along a major security

concern. Wireless communication is exposed to diverse varieties of attacks, including Denial of Service attacks, node cloning, node capture, physical tempering and number of other attacks. Since these nodes are physically limited in size and resources, implementing a secure network is among the fundamental research challenges especially when WSN is gaining a rapid influence in industry, academics and defense [2].

Among the application areas of WSN, one of the leading application area is combat zone monitoring. Such application areas require security to be of paramount importance. In such condition integrity, confidentiality, authentication, availability, freshness and scalability of network are very significant tasks. If these issues are not properly handled, they can result in significant security breaches. Which puts a question mark on data reliability [3]. Among the number of capabilities of WSN is its ability to self-organize its network with complete coordination and corporation among the nodes [2]. Leaving security parameters, a much more difficult and highly significant task for the researchers. Regular public cryptography

¹ Computer Science Department, Sukkur IBA University, Sukkur, Pakistan

² Computer Science Department, SZABIST Larkana Campus, Larkana, Pakistan

Corresponding Email: Murtaza.siddiqi@iba-suk.edu.pk

approaches or techniques were designed by not keeping in view the resources limits. That is why traditional cryptography or security methods are not considered very suitable for WSN security implementation.

This paper is divided into 8 sections, so that each information relevant to WSN security can be covered comprehensively. Introduction is covered in section 1. Section 2 covers the security goals of WSN. Section 3 describes the challenges of WSN due to which security is a challenging task in WSN and section 4 covers a general network architectural of WSN. In Section 5, some of the well-known attacks in WSN are being discussed. Section 6 contain the analysis on some of the purposed solutions from different researchers and section 7 contains the conclusion of the paper.

2. Security Goals

Some of the security goals of WSN are similar to that of a highly distributed database. Since security related goals for distributed database are already thoroughly researched and implemented by researchers. Which can be summarized as: Data only accessible to authentic users (to achieve confidentiality), data should be authentic or genuine (to achieve integrity), availability of Data to authentic user.

The above-mentioned security goals are also applied on WSN. As from a user point of view, WSN and distributed database is a single entity. To understand the security goals in a much better way, security goals can be classified as outside security and inside security [4]. The outside security for distributed data base were mentioned earlier at the beginning of section 2, as for WSN. The outside security goals are much clear as query processing [5], access control [6] and large scale anti-jamming services [7].

As far as inside security goals are concerned, WSN differs from distributed database system. Inside security parameters for WSN can be stated as resilient, confidential, scalable and authenticity while

communicating between nodes [8]. These mentioned inside security parameters also include a number of tasks which WSN performs internally, which again can be categorized as within network processing, data aggregation, routing, data storing (within node or sink). Apart from the mentioned inside and outside security parameters, WSN also contains a number of other challenges. These challenges are discussed in section 3 of the paper.

3. Challenges of WSN

WSN exhibit unique nature and have range of challenges that must be considered when addressing security concerns. Security goals cannot be achieved without understanding the challenges, which come along WSN.

3.1. Customization

Due to unique attributes and characters of WSN almost every aspect of the device has to be considered and customized. Software and hardware requirements of WSN are quite different and so are the requirements of operation.

3.2. Resource limitations

Traditional security mechanisms require high resources as they have high overheads that are not suitable for WSN, keeping in view the resource limitation. Many security approaches are computationally expensive, thereby leading to energy overheads [9].

3.3. Absence of Central Control

It is often challenging to have a central point of authority in WSN, because of their enormous scale, resource limitations, and network deployment. Consequently, security solutions must be dispersed and nodes must cooperate to accomplish security. As node related issues are very common with WSN [10].

3.4. Isolated Location

The most important step to provide security is to offer precise, authentic and controlled physical access to a sensor node. Many WSN are left unattended, because they are operated in remote and hard-to-reach sites, as they are deployed in open environments. So constant monitoring and physical protection to a sensor node is very difficult, making it vulnerable to unauthorized physical access. Nodes, which are physically tampered and are being compromised can later on cause a number of security breaches [9].

3.5. Error-prone communication

Packets being exchanged between nodes or sink in WSNs might be corrupted or even lost due to a variety of causes, including channel errors, routing failures or collisions. Such packet related issues can affect security mechanisms or overall operational ability of a network [9].

3.6. Scalability

As sensor nodes are prone to failure, deployment of new nodes is necessary. Nodes that are being deployed must be able to quickly authenticate and be part of the network operations. With these new nodes becoming part of the network, the network must also be equipping with mechanisms for swift and rapid authentication and the ability to adapt with the changing topology [10].

3.7. Hardware constraint

Due to limitation in size, the hardware for WSN has to be specialized. With this limitation and low cost factor, WSN nodes need to be rigid and in case of faulty node the coordinator or the sink node should be able to detect it immediately [10].

3.8. Energy Constraint

The most fundamental or major issue in WSN is power management and watchful use of existing energy. Approaches for energy management in sensor networks can be generally separated into two groups: active and passive methods. Active methods to save

energy comprise focused operating systems like watchdog timers, using sleep states or using flexible voltage processing. Passive methods comprise sophisticated energy sources to replace the batteries and positioning of sensors into power efficient topologies [9, 10].

3.9. Time Synchronization

Until now, a “flawless” alternative for the time synchronization concern in sensor networks has not been established. Several of the concepts that are used for time synchronization can be categorized as explicit synchronization and peer-to-peer synchronization. In explicit synchronization clocks are not kept synchronized at all time, instead in order to put less load on the communication overhead, every node retains its own individual timescale. Therefore, exchange of information between dissimilar time scales is carried out “on demand”. On the other hand, peer-to-peer synchronization clocks are simply maintained synchronized between neighboring nodes. The rationale explanation for this concept is that communication among neighboring nodes includes only those nodes that are synchronized [9].

3.10. WSN Node

WSN node plays a very important part in security implementation, as a node is very small and limited in its resources. A WSN node is usually equipped with one or more sensors, a wireless transceiver for communication (i.e. antenna), a microcontroller and memory module. While software component for a node may include specially designed operating systems (i.e. TinyOS, LiteOS) to full fill WSN node’s specific requirements with in the provided resources. Such Operating system scan process received data, acquire sensed data from the sensors; organize sensed data for transmission, resources management and basic network related tasks [10].

4. General Architecture

After discussing the challenges in WSN, section 4 covers the general architecture of WSN. As it is important to have an understanding of network architecture to properly understand the challenges of a network. There WSN architecture is of two types, hierarchical and flat. Based on architecture the sensors organize themselves to achieve specific goals. In flat formations, all the nodes participate in the decision-making procedure and play equal part in internal routing protocols. On the other hand, in hierarchical arrangement the network is separated into clusters or group of nodes. A single unit called "cluster head" makes organizational decisions, like data aggregation. It must be observed that it is as well probable to have a mixture of the two previous formations into the same network; for example, to avoid circumstances where the "spinal cord" of the network or the cluster heads fails then the information must be routed to the base station by alternate means. This can be achieved by using a combination of both architectures [11].

4.1. WSN topologies

In general, WSN are organized in three basic type of topologies star, cluster and mash [12]. With recent development and wide area of application new topologies are also being introduced in WSN. Topologies that already exist in traditional networks including tree, ring, circular and grid are also utilized in WSN [13]. Among these general topologies grid topology is energy efficient in theoretical comparison [13].

Next section will cover the most common attacks that are being inflicted on WSN.

5. WSN Attacks

In general, we can classify the attacks as active and passive. Passive attacks are attacks that involve eavesdropping or information gathering without raising any red flags. On the other hand, Active attacks are more aggressive and are highly contagious to the network.

Active attacks can involve modification or destroying a packet, providing wrong routes to the network or even jamming the network. In some literature, active and passive attacks are categorized under goal-oriented attacks. The other type of attack category tries to deteriorate the performance of the network such attacks are commonly called performance oriented attacks. Performance oriented attacks can be conducted from within the network and from outside the network and then there are attacks which are layers based attacks [9].

Layer based attack exploit vulnerabilities at different layers (physical, data, network, transport and application) to cause harm to the network. Numerous categories and patterns of attacks on WSN are being documented and discussed in different literatures. In this section, the most known and common types of WSN attacks are being discussed.

5.1. Eavesdropping

An attacker with powerful resources can passively gather or collect information from the WSN in case the network is not well protected in terms of encrypted during communication between nodes or sink [14].

5.2. Node based attacks

If a node is physically accessed or captured by an attacker, then the attacker can conduct a number of attacks that may include black hole attack, Sybil attack, wormhole attack, clone attack [15]. As a node can reveal information that can be very useful to the attacker, information including cryptographic keys, network architecture or node ID thus compromising the entire network. Attacker can also use such information to deploy a false node. False node can insert malicious data or if it is robust enough it can even decoy other nodes to send data to it.

Then there are always possibilities that the node malfunctioned, resulting in generating inaccurate data. If that malfunctioned node is a cluster head, then a much worse condition can be expected. In case of a cluster head

outage or malfunction, robust protocols must be implemented in order to nominate a new cluster head and continue with network operations without much wastage of time and resources [14].

5.3. Attacks based on traffic analysis

There is always a possibility that is based on communication patterns and sensor activity analysis; an attacker can acquire enough information to organize a well versatile attack. Despite the encrypted communication, attack based on such analysis can create security issues in WSN [14].

5.4. Sybil Attack

Sybil attack can be defined as, when a single node presents numerous identities to network nodes. Such node is a substantial threat to routing protocols, especially when location aware routing is a requirement for nodes to coordinate and exchange information with their neighbors for efficient geographical routing. Normally a networks authentication mechanism or sequential analysis can avoid or detect a Sybil attack from an outsider [15, 16].

5.5. Sinkhole/Black hole attacks

In a sinkhole/blackhole attack, the aim of the attacker is to lure the traffic from neighboring nodes to its compromised node. Compromised node act as a sinkhole or a black hole and drops all the traffic or packets it receives from the network [17, 18].

5.6. Jamming

Jamming is a well-known attack in wireless communication. The main idea behind such attacks is to disturb the radio channel by sending information on the frequency band being utilized by the targeted network. *Denial of Service* comes under the types of attack that uses jamming technique to disrupt communication. While jamming is usually conducted at physical layer, denial of service attacks is normally conducted on data link layer [19, 18].

5.7. Exhaustion

Such attacks target the resources of a node by forcing node in to performing operations, which are simply not required, and result in waste of time and energy [20, 21].

6. Analysis on Attacks and Counter Measures

In section 6 analysis are performed on the counter measures, which are being developed by different researchers against some of the most common WSN attacks.

6.1. Summary of analysis

As per analysis, the methods purposed by most of the researchers are very comprehensive and are as per requirements of WSN security. Nevertheless, the additional resource requirements to implement such methods, is on the higher side. Some purposed methods require more processing power, while some follow a lengthy process of authentication. Such lengthy process could result in extra time consumption and can affect the freshness of data. While location based routing algorithms require additional mechanism for geographical positioning. After the analysis in table 1, it can be stated that much work is still required to achieve a comprehensive security suite for WSN.

7. Conclusion

The most concerning issue when considering WSN security is the unattended environment, random deployment, node size and limited resources. Due to node size, resources are very limited, making it very difficult for security experts to design a rigid mechanism with in the provided resources. Most of the approaches discussed and purposed in literature may be implemented in selective WSN models, but cannot full fill the requirement of general WSN models. So the hunt for more effective, smart and rigid security methods for WSN continue for researchers in coming days [9, 20].

TABLE 1. Analysis on different Counter Measures.

Attack	Counter Measure	Analysis
Jamming, Node Tampering and Eavesdropping.	A variety of traditional attacks that target wireless medium can be countered by Spread-Spectrum based techniques. Other counter methods may include Tamper proofing, enhanced key management schemes and encryption. In some cases, directional antenna access for access management can also be used [22]. Coalesced neighbour nodes can be used to avoid jamming regions [7].	Mentioned counter measures require additional resources and enhancement in security mechanism. Keeping in view WSN resource constrain it's very difficult to acquire resources for such enhanced measures, rest aside make physical enhancements which could result in enlarging the size of the node. While if geographical counter measures are used, they might open new security vulnerabilities which come along with geographical routing algorithms.
Exhausting, generating malicious traffic to overcrowd or jam communication channel i.e hello flood attack	To counter such attacks Spread-Spectrum based techniques can be utilized. Other than that algorithm can be implemented, which can limit the data rate or can black list a node using MAC, which generates unusual traffic patterns [22]. Data forwarded by nodes can be checked for false information and such information can be dropped and should not be forwarded [23]. Probabilistic based sharing of secrets, bidirectional verification and routing based on multi-path multi-base station [24]. Use link layer to strengthen data integrity and message authentication [8].	Documented techniques provide a comprehensive counter against the mentioned attack categories. However, implementing such measures will again require additional work at node level, resulting in additional resource requirements. Even if such counter methods are implemented at sink node or cluster head (in case of hierarchy architecture) those nodes will have to perform aggregation, authentication (data and new node), filtering of data, forwarding and other similar tasks, which will not only be an overburden but will also result in lack of performance at operational level of the network. We must also keep in mind that freshness of data also plays

		<p>a vital role in WSN. Methods that require multidirectional or multi-level verification can result in delay or extra time consumption.</p>
<p>Sybil Attack, Sinkhole, Wormhole, false routing information</p>	<p>Mentioned category of attacks can be countered using flexible routing algorithms, multi direction authentication and handshake mechanism, monitoring of traffic, restriction to routing access, invalid route detection and reporting methods [22]. Registration process, pre-distribution of random key, geographical position verification, to detect a Sybil entity a code attestation with local based verification [25]. Use of Temporal leases, time synchronization within network or all the communication devices and symmetric cryptography [26]. Using broadcast inter-radio behaviour to observe neighbour transmissions and to detect any suspicious activity similar to black hole attack, use geographical routing algorithm [27].</p>	<p>Most of the counter measures against mentioned attacks focus on two ideas. One geographical location based routing mechanism and second time synchronized based mechanisms. With these counter approaches WSN needs additional time synchronizing mechanism and keeping in view such methods will require very precise and calculated mechanism. Including timely verification method, so that communication devices should always be synchronized. As for geographical location based routing requires some kind of ability within the node to detect location and coordination with surrounding nodes with the help of location based routing algorithms. Most of the location based algorithms need broadcasting at initial authentication that can be a concern from security point of view.</p>

REFERENCES

- [1] O. Althobaiti, M. Al-Rodhaan, and A. Al-Dhelaan, "An Efficient Biometric Authentication Protocol for Wireless Sensor Networks," *International Journal of Distributed Sensor Networks*, v. 2013, Article ID 407971, 13 pages.
- [2] M. A. Khan, G. A. Shah, "Muhammad Sher "Challenges for Security in Wireless sensor Networks (WSNs)," *International Journal of Computer and Information Engineering*, vol. 5, no. 8, 2011.
- [3] G. V. Crosby, L. Hester, and N. Pissinou "Location-aware, Trust-based Detection and Isolation of Compromised Nodes in Wireless Sensor Networks," *International Journal of Network Security*, vol.12, no. 2, pp. 128–138, March 2011.
- [4] Z. BENENSON, P. M. CHOLEWINSKI, and Felix C. "Vulnerabilities and Attacks in Wireless Sensor Networks," *FREILING.Laboratory for Dependable Distributed Systems, University of Mannheim, 68131 Mannheim, Germany* SAP - Research and Breakthrough Innovation, Germany. Wireless Sensor Network Security, J. Lopez and J. Zhou (Eds.) IOS Press, 2008.
- [5] L. Hu and D. Evans, "Secure aggregation for wireless networks," *In SAINT-W'03: Proceedings of the 2003 Symposium on Applications and the Internet Workshops (SAINT'03 Workshops)*, page 384. IEEE Computer Society.
- [6] H. A. Maw, H. Xiao, B. Christianson, and J. A. Malcolm, "A Survey of Access Control Models in Wireless Sensor Networks," *Journal of Sensor and Actuator Networks*, vol. 3, pp. 150-180, 2014.
- [7] A. D. Wood, J.A. Stankovic, and S. H. Son, "JAM: A Jammed-Area Mapping Service for Sensor Networks," *24th IEEE Real-Time Systems Symposium, RTSS, 2003*, pp. 286-297.
- [8] C. Karlof, N. Sastry, and D. Wagner, "TinySec: A link layer security architecture for wireless sensor networks," *In Second ACM Conference on Embedded Networked Sensor Systems (SensSys 2004)*, November 2004.
- [9] K. Chelli, "Security Issues in Wireless Sensor Networks: Attacks and Countermeasures," *Proceedings of the World Congress on Engineering 2015*, vol. I WCE 2015, July 1-3, 2015, London, U.K.
- [10] S. Patil, V. Kumar B P, S. Singha, and R. Jamil, "A Survey on Authentication techniques for Wireless Sensor Networks," *International Journal of Applied Research*, ISSN 0973-4562, vol. 7, no.11, 2012.
- [11] A. Davis, H. Chang, "A survey of wireless sensor network architectures," *International Journal of Computer Science and Engineering Survey (IJCSES)*, vol. 3, no. 6, December 2012.
- [12] M. Rajput, U. Ghawte, "Security Challenges in Wireless Sensor Networks," *International Journal of Computer Applications*, vol. 168, no. 5, June 2017.
- [13] D. Sharma, S. Verma, and K. Sharma "Network Topologies in Wireless Sensor Networks: A Review," *IJECT*, vol. 4, no. Spl - 3, April - June 2013.
- [14] T. Zia and A. Zomaya, "A security Framework for Wireless Sensor Networks," *IEEE Applications Symposium, Houston, Texas USA*, February 2006.
- [15] Abirami. K, Santhi. B, "Sybil attack in Wireless Sensor Network," *International Journal of Engineering and Technology (IJET)*, vol. 5, no. 2, Apr - May 2013.
- [16] P. Raghu Vamsi and K. Kant, "Detecting Sybil Attacks In Wireless Sensor Networks Using Sequential Analysis," *International Journal On Smart Sensing And Intelligent Systems*, vol. 9, no. 2, JUNE 2016.

- [17] I. Raju and P. Parwekar, "Detection of Sinkhole Attack in Wireless Sensor Network," *Proceedings of the Second International Conference on Computer and Communication Technologies. Advances in Intelligent Systems and Computing*, vol. 381, Springer, New Delhi, 2016.
- [18] F. Hu and N. K. Sharma, "Security considerations in ad hoc sensor networks," *Ad Hoc Networks, Published by Elsevier Science*, 2005, pp.69–89.
- [19] F. Anjum and S. Sarkar, "Mobile, Wireless, And Sensor Networks Technology, Applications, And Future Directions," *IEEE Press*, 2006.
- [20] M. L. Messai, "Classification of Attacks in Wireless Sensor Networks," *International Congress on Telecommunication and Application'14 University of A. MIRA Bejaia, Algeria*, 23-24 April, 2014.
- [21] P. Adrian, J. Stankovic, and D. Wagner. "Security in Wireless Sensor Networks," *Communications of the ACM*, vol. 47, pp. 53-57, 2004.
- [22] D.G. Anand, Dr.H. G. Chandrakanth, Dr. M. Giriprasad, "SECURITY THREATS & ISSUES IN WIRELESS SENSOR NETWORKS," *International Journal of Engineering Research and Applications (IJERA)*, vol. 2, no. 1, pp.911-916, 2012.
- [23] F. Ye, H. Luo, S. Lu, and L. Zhang, "Statistical en-route filtering of injected false data in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 839 – 850, April 2005.
- [24] M. A. Hamid, M-O. Rashid, and C. S. Hong, "Routing Security in Sensor Network: Hello Flood Attack and Defense," *to appear in IEEE ICNEWS 2006*, 2-4 January, Dhaka.
- [25] J. Newsome, E. Shi, D. Song, and A. Perrig, "The sybil attack in sensor networks: analysis & defenses," *Proc. of the third international symposium on Information processing in sensor networks, ACM*, 2004, pp. 259 – 268.
- [26] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Packet leashes: a defense against wormhole attacks in wireless networks," *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE INFOCOM 2003*, vol. 3, 30 March - 3 April 2003, pp. 1976–1986.
- [27] Z. Karakehayov, "Using REWARD to detect team black-hole attacks in wireless sensor networks," in *Workshop on Real-World Wireless Sensor Networks (REALWSN'05)*, 20-21 June, 2005, Stockholm, Sweden.

Swarm Based Coverage Using Multiple Informed Leaders

Ahmad Din¹, Ashfaq Ahmed¹, Kashif Zia¹, Abbas Khalid¹, Owais Khan¹

Abstract:

Robotic exploration typically involves navigation through unknown terrains. In this paper, collaborative strategy for coverage is presented, in which the concept of informed agents as a leader in a swarm has been introduced for autonomous coverage. The small population of robots in large swarm act as informed leaders and help others to accomplish their tasks related to the exploration. These informed leaders receive information about the environment from external sources (e.g. humans, media etc.), and influence uninformed robots using their swarm behaviors. Multiple Swarm behaviors have been designed for swarm navigation, and dynamic selection of the informed leaders. This approach has been tested in simulation of homogeneous swarm, with and without informed leaders. The number of informed agents needed to guide the swarm effectively has been investigated as well. Experimental results showed that the introduction of informed agents improves the coverage task.

Keywords: *Swarm Intelligence; Swarm robotics; Multi robotic exploration; leader selection.*

1. Introduction

Autonomous robotic exploration and coverage has been a hot research topic in which a robot visits the environment, build the map, and localize itself within the map simultaneously. If an environment is dangerous, remote, or expensive for human access, a mobile robot may help to build a map of its surroundings and navigate based on the map [1]. Furthermore, robots are resource-constrained specially in multi-robotic environments (where robots collectively perform a task) in which coverage is one of the key issues to be addressed effectively. Coverage problem has applications in exploration, navigation, search and rescue operations, particularly at city scale. In a multi-robotic environment, coverage becomes an issue due to limited capabilities of the robots, in terms of memory and computation. Either this can be achieved in a completely distributed fashion, where robots can only sense each other and do not communicate [2] or it can be achieved where robots can sense as well as communicate with each other. The behavior based coverage and exploration

using multi-robots presented in [2] has been extended for the swarm. We have focused on leader based swarm coverage. Dynamic Leader selection is one of the challenges that have been addressed in this research. In this regard, social science based leader selection mechanisms have been considered as a starting point [3]-[4]. We have refined the model presented in [4], so that it works for large scale coverage. In addition to mobility issues, such as target selection, obstacle and collision avoidance, and density-based speed, we have also focused on natural group-based mobility paradigm; the swarm robotics [5]. The informed leaders are guiding other robots in the swarm to maximize the coverage rate. We have examined the usability of swarm base mobility as compared to the individual decision making using GIS map based simulations.

2. Related Work

Rekleitis et.al [18] proposed an exploration algorithm, in which one robot explores environment, while two stationary robots observe it, making a triangle shaped group [6].

¹ Department of Computer Science, COMSATS Institute of Information Technology, Abbottabad
Corresponding Email: ahmaddin@ciit.net.pk

The main disadvantage of this techniques is a centralized control. In a later work, the same authors presented a technique in which couple of robots track the environment to search each other, for better mutual localization [7]. Arkin et.al suggested a behavior based exploration [8], where the “wander” behavior acts as communication actor, and “informed exploration” behavior helps to use map to explore the given arena. Powers et. al used “motor-schemas” [9] to preserve line-of-sight communication among members of a team [10]. This technique focuses on value based approach, in which all robots agree on a direction, then they move toward that direction, they make decisions in a distributed way, but stay connected every-time. Nguyen et al. [11] developed an actual real-world system using leader-follower approach. A similar leader-follow model has been designed by Howard and his team members [12] in which eighty robots were used to explore the building, and transmit all the acquired information to the remote operator. Yamauchi presented a revolutionary work in 1998 in which new exploration strategy using the concept of the “frontiers” [13] was introduced. In this technique, frontiers are the borders between visited area and un-visited area. In addition of Yamauchi’s frontier-based exploration, Simmons et.al proposed a technique in which robots used ‘bids’ based on estimates of the travelling cost to major locations, and information gain [14]. These bids were submitted to chief agent, which assesses and allocates tasks based on maximum utilization. Stachniss et al. used place labels to define structures of the environment that can be used instead of frontiers based approach [15]. Wurm et.al’s [16] work classifies unvisited area into segments, which are computed by Voronoi graph. Koenig et.al studied ant patterns and developed ant-like robot to explore [17]. Our approach is also inspired from nature and is using behavior-based approach. The swarm is steered toward the unexplored parts of the environment using multiple leaders.

3. Behaviour Based Swarm Design

3.1. Environment

To simulate this problem, a portion of Vienna city map is used, which contains parks, roads, streets, narrow streets. The shape files of this GIS maps are imported into NetLogo, and this map is segmented into the small grids called cells. Furthermore, walkable, and non-walk-able cells were defined. Fig. 1 shows the screenshot of the whole environment.

For ease and visibility, walkable portion is shown in yellow, buildings in brown and robots in red color. In this environment, obstacles can be placed at any part of the map beside predefined buildings and other obstacles.

3.2. Behaviors

In this approach, it is assumed that robots have limited processing and memory. Therefore, behavioral based architecture is used for collective mobility and control of the swarm. Many behaviors have been designed including obstacle avoidance, avoid visiting past locations, Locating open area, leader selection, etc.



Fig. 1. GIS based environment for Robotics Coverage Simulation.

3.2.1. Obstacle Avoidance

It is the basic behavior in which robot detects the obstacle using local sensing capabilities and avoids these obstacles. This behavior is fused with other behaviors and it

avoids obstacles in such a manner that it could get largest free space to navigate.

3.2.2. Avoid Visited Locations

This behavior helps robots to navigate toward the newest locations. For this behavior, relatively long-range sensors are used as compared to the obstacle avoidance behavior. It keeps the track of previously visited locations. We have used efficient data structure to store and retrieve the information about the visited locations. Whenever robot coordinates are changed, it checks if cells ahead is already visited. If any of the cell up to predefined length is already visited, then it will rotate 45 degrees and check again for the patches ahead. It keeps on doing the same for all 8 possible angles until it returns to the same position. If all are visited, it moves to the patch ahead.

3.2.3. Locate Open Space

This behavior locates a wide area in which robot can search for targets, this may be an open room or hall. In our case, we have used this to locate open streets, roads, and open grounds like parks. It is a very useful behavior since it saves time going to closed type rooms and round shaped areas. In multi-robotic autonomous exploration, usually robots stuck in the corners and wasting time over there. This behavior also addresses this issue by gathering information of the largest free space around the robot's current location. Whenever this behavior is triggered then robots stops its movement and starts searching open space around itself.

3.2.4. Leader Selection

This is most important behavior regarding our contribution in this work. We have tested many leader selection algorithms to find which best fits in our case for search and rescue of large open space area. We studied motion and movement of ants and birds and their swarm principles then we suggested our strategy which is more near to the nature. Moreover, our strategy for leader selection

(informed agent) is dynamic i.e. leader changes with the passage of time based on the shape of swarm. As the direction of swarm changes then leader is changed if it is necessary. Leader selection algorithm runs after every 30 ticks and elects for new leader if necessary. When leaders receive goals then they start moving toward their goal. Other agents observe the leader nearby them which helps swarm to move toward the goal. The algorithm of the leader selection is shown in Algorithm 1.

3.3. Inter-Agents Communication

In our approach, robots communicate the speed, previously visited locations, and other information with agents in their neighborhood. In this implementation, we are using Stigmergy, which is indirect communication mechanism. This technique is inspired from ant communication. Though in future, direct and explicit communication techniques can be used once this algorithm is extended to the real robots.

Algorithm 1. Leader Selection.

Procedure SETUP

```

if ticks > maximumTicks then
  set percent = 10;
  set ticks = 0;
  set isLeader = false;
  set neighbor_agents =  $\varnothing$ ;
  broadcast message "No leader";
  SELECTLEADER(percent);
end if
end procedure

```

Procedure SELECTLEADER (int PERCENT)

```

set neighbor_agents = ask all agents to broadcast
their IDs, and no of other agents in radius of their
vision;
sort neighbor_agents;
max_leaders = (total_agents * percent)/100;
ask neighbor_agents[max_leaders] top entries to
set isLeader = true;
ask leaders to broadcast message "I am leader";
broadcast message "Follow new leaders";
set ticks=ticks + 1;
end procedure

```

3.4. Informed Agent and Swarm Control

Informed agents receive information about the unexplored region which is considered as goal. All the leaders receive the goal information simultaneously, so that swarm could be steered toward the goal. In real world situation, the goal information about unexplored and important regions can be received from other sources, e.g. human operator, media, other robots like aerial robots. Beside the goal information, global heading and speed is computed for leaders. Other robots move based on the local control laws of the swarm. These local control laws include align, attract, and repel. These control laws are fused with other behaviors described earlier including obstacle avoidance, avoid visited locations etc. Alignment and speed of the normal robots (robots other than leader) is the average of the other robots in the neighborhood, but if there is leader in the neighborhood, the heading and speed of the leader is given higher weightage, which helps the robot to be aligned with leader. This is how leaders influence the alignment and speed of the swarm. Robots avoid collisions and cohesion in swarm using repulsion and attraction behaviors. Attraction, repulsion, and alignment are fused with other behaviors, and normalized to compute the motion of the swarm. Leader in the swarm is selected dynamical based on leader selection algorithm.

3.5. Coverage Calculation

In this work our focus is to explore the whole area, we have used the formula derived for coverage as described by Schwager et.al. [19]. The coverage is the percentage of area explored by all the robots in a swarm. In our case, total walkable patches are counted, and the total patches visited by the swarm. So, we can easily estimate coverage in percentage using the equation defined below.

$$Coverage = \frac{AreaCovered}{TotalTargetArea}$$

4. Experiments, Results and Discussion

To validate our strategy using informed leaders, dynamic selection of leader, goal generation, and its effects on the coverage, series of different experiments were conducted. Each experiment was conducted 3 times to get most optimal reading. To simulate our proposed strategy, we used NetLogo simulator. Moreover, all simulations were carried out on computer having following specification.

- Intel Core i3, 3rd Generation Processors.
- Intel 4000 HD Graphics card.
- 8 GB of RAM.
- Linux Mint operating System.

We have compared proposed technique with the swarm navigates in the environment using random walk. Secondly, we have evaluated the effect of number of the informed leaders. These leaders receive the clues about major goals in the environment. We used 10% and 20% leaders of the entire population [19]. When they all move toward their goal, it creates a force toward some major points of environment, which is easily observed by the other agents, the mechanism is explained in the previous section. As goal information is communicated to the swarm, it drives the swarm toward the unexplored regions. Therefore, full coverage of the environment is possible. The stopping criteria for the experiment is full coverage of the environment for both with and without informed leaders' cases.

4.1. Agents Placement Technique

We have tested two types of agent placement strategies, "Randomly Distributed" and other one is "Group distributed".

In randomly distributed, we distributed all the population over the whole area randomly, while in group distributed strategy we make one or two groups of all the available population. Size of group is entirely random.

Experiments are characterized according to these placement strategies. Total 7 different experiments were conducted. In experiment 1-

3 Randomly distrusted strategy were used, while in experiment 4-7 group distributed strategy is used. Detailed information about these experiments can be seen in table 1. Each experiment is conducted 3 times.

TABLE 1. Experiments details.

Strategy	Experiment	Agents	Leader %
Randomly Distributed	1	100	0
	2	100	10
	3	100	10
Group Distributed	4	100	0
	5	100	10
	6	100	10
	7	100	20

4.2. Experiment 1

First experiment was conducted with total of 100 agents which were randomly distributed over the whole environment. Agents can re-visit the visited places. Table 2 shows average and Standard deviation of three combined runs. Figure 2 shows exploration graph w.r.t time (in minutes).

TABLE 2. Average and SD for Experiment 1.

Run	Average	Standard Deviation
3	565 min	47.08 mins

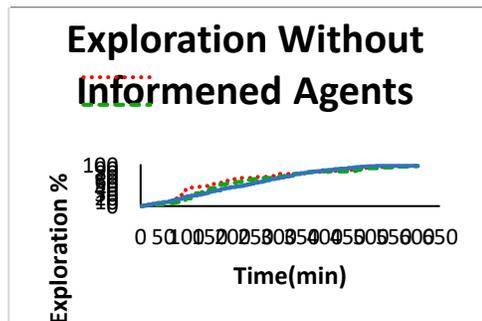


Fig. 2. Three Combined runs of Experiment 1.

Almost all the runs took around 10hours. We can see a lot of flat zones (in Figure. 2), because the exploration was conducted without any informed agents and also by avoid visited location behavior. This let the agents to visit the already visited locations. We simply call this a redundancy.

4.3. Experiment 2

Experiment 2 was conducted with the same configuration as for the previous experiment except now we have 10% informed agents. Average and Standard deviation of combined three runs can be seen in table 3. Figure 3 shows graph of three runs w.r.t time(in minutes).

TABLE 2. Average and SD for Experiment 2.

Run	Average	Standard Deviation
3	205.3 min	25.23 min

In figure 3 we can see that our flat zones were reduced. This is because we used 10% of informed agents but still we are lacking avoid visited location behavior that still costs us redundancy. Informed agents know some places of the environment that are still unvisited so they take the group to that places which leads us a quality exploration in corresponding against previous experiment. It is clear that by introducing informed agents total time taken to explore the whole environment drops to almost a half.

4.4. Experiment 3

Third experiment was conducted with 10% informed agents and with avoid visited locations behavior. Average and Standard deviation is listed under table 4. Figure 4 shows graph of exploration w.r.t time.

TABLE 3. Average and SD for Experiment 3.

Run	Average	Standard Deviation
3	205.3 min	25.23 min

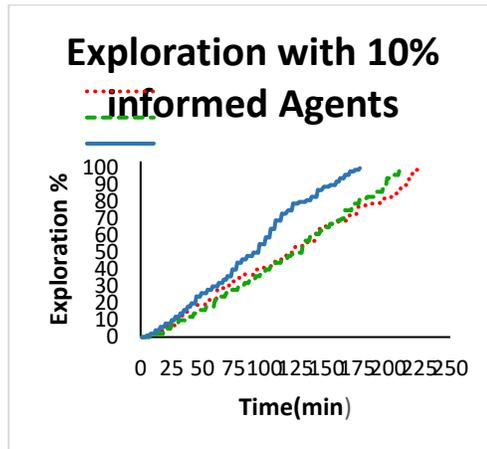


Fig. 3. Three Combined runs of Experiment 2.

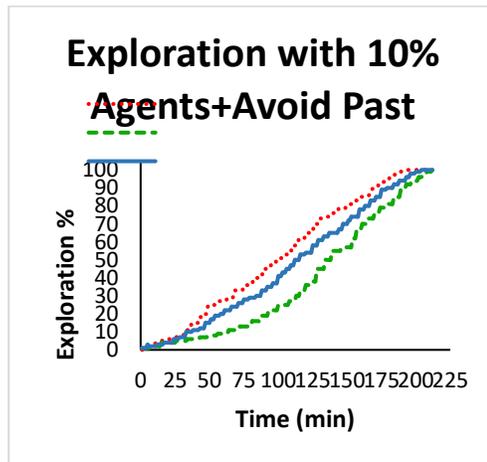


Fig. 4. Three Combined runs of Experiment 3.

By looking at average of this experiment we do not see any major progress by introducing informed agents but still we improve our standard deviation. Also, this strategy was randomly distributed so agents wasted most of the time in making swarm.

Avoid visited location saves us some time but still we need improvement in time. Figure 4 also shows that we have very low flat zones.

4.5. Experiment 4

Experiment 4 was conducted with group distributed strategy. In this experiment, we used single group with no informed agents. All the agent population was placed in a single group and were randomly assigned to some walkable portion. Table 5 shows average and standard deviation of all three runs. Figure 5 shows graph of exploration v/s time.

TABLE 4. Average and SD for Experiment 4.

Runs	Average	Standard Deviation
3	270 min	12.5 min

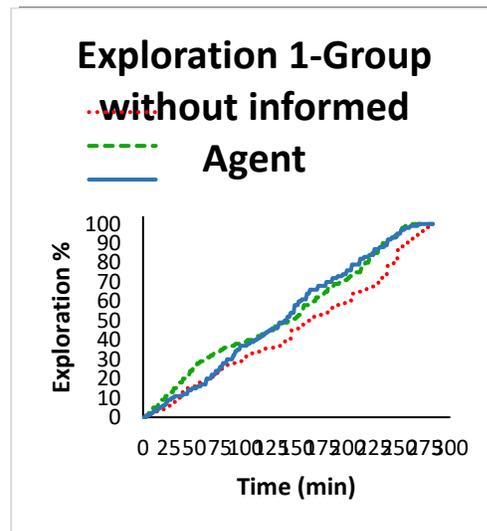


Fig. 5. Three Combined runs of Experiment 4.

As the exploration started all the single group swarm started to move in one direction but as the time passes it broke into several groups. As we are lacking informed agents and avoid visited location, behavior agents started to wander which led them to visit already visited location. By comparing it to previous experiment, total time taken to complete the exploration increased.

4.6. Experiment 5

Experiment 5 was the same as the experiment 4 but this time we used 10% leaders (informed agents) over the entire population. Average time and Standard deviation of combined three experiments can be seen in table 6. Figure 6 shows graph for all three runs of exploration v/s time.

TABLE 6. Average and SD for Experiment 5.

Run	Average	Standard Deviation
3	132.3 min	12.5 min

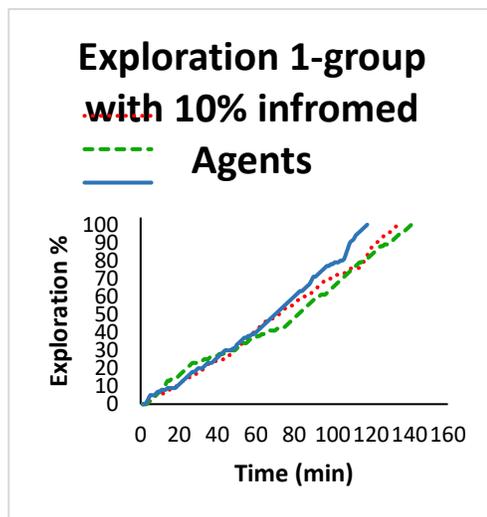


Fig. 6. Three Combined runs of Experiment 5.

It is clear to see in Table 6 that total average falls up to 50% if we compare it with experiment 4. By enabling informed agent behavior, the single group broke into several small groups. Informed agents took all the agents in their swarm to major location which led to a decent exploration.

4.7. Experiment 6

The only Change made in current experiment was to enable avoid visited location behavior. Figure 7 shows graph of whole three exploration v/s time, and Table 7

shows average and Standard deviation of all three combined runs performed for this experiment.

TABLE 7. Average and SD for experiment 6.

Run	Average	Standard Deviation
3	81.6 min	2.0 min

It is very clear to see that in Table 7 by using 2-group strategy with avoid visited locations and informed agent behavior our average is improved. Total time taken to cover the full environment fall up to around 1.5hours. The exploration was very smooth and informed agents led the groups to unvisited locations.

4.8. Experiment 7

In this last experiment, we used 20% informed agent rather than 10%, in-order to see whether increasing informed agents will effect the overall performance or not. Rest of the configuration was the same as in the experiment 6. Table 8 shows average and standard deviation of all the three runs which were performed for above mentioned experiments.

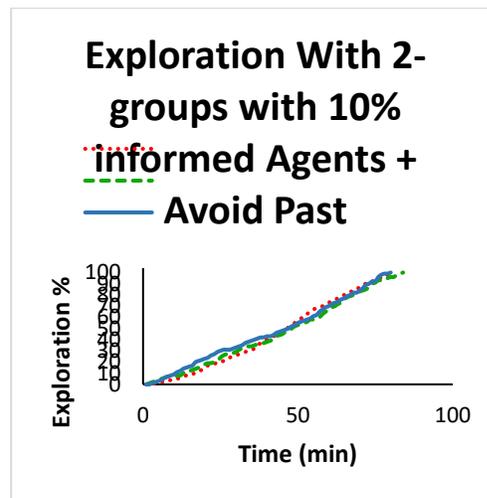
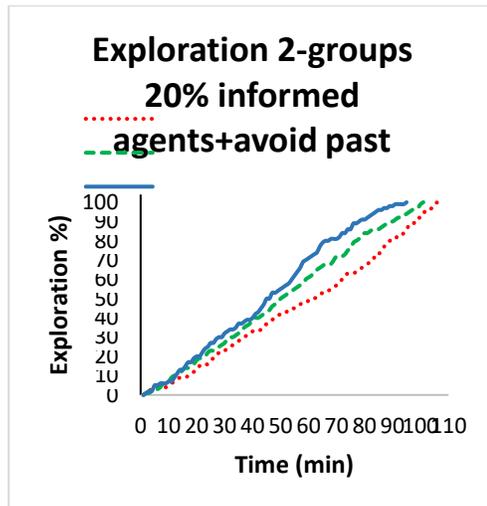


Fig. 7. Three Combined runs of Experiment 6.

TABLE 5. Average and SD for Experiment 7.

Run	Average	Standard Deviation
3	100.6 min	5.05 min

**Fig. 8.** Three Combined runs of Experiment 7.

In this last experiment, we used the same configuration as in the experiment 6. The only change was in the percentage of informed agents. In order to see influence of informed agent we used 20% instead of 10%. As discussed earlier, informed agent does not get whole map of the environment. They just got clues about major goals. As major goals are limited and not all informed agents got unique goals, a single goal was shared between more than one informed agent which led multiple groups (swarms) to a single location. Hence overall exploration effected.

5. Conclusion

In this paper, we have suggested a coverage technique for large scale environment. The basic implementation of this type is system is feasible in a disaster or search and rescue operation. As our results show that in large scale environment informed agents played key role in exploration of an unknown area. Also,

it was a behavior based technique so agents decide which behavior to call in which type of situation.

REFERENCES

- [1] S. Mac, et al. "From theory to practice: Distributed coverage control experiments with groups of robots," *Experimental Robotics*. Springer, Berlin, Heidelberg, 2009.
- [2] J. S. Cepeda, L. Chaimowicz, R. Soto, J. L. Gordillo, E. A. Alanís Reyes, and L. C. Carrillo-Arce, "A behavior-based strategy for single and multi-robot autonomous exploration," *Sensors*, vol. 12, pp. 12772--12797, 2012.
- [3] S. Wu and Q. Sun, "Computer simulation of leadership, consensus decision making and collective behaviour in humans," *PloS one*, vol. 9, 2014.
- [4] J. R. Dyer, A. Johansson, D. Helbing, I. D. Couzin, and J. Krause, "Leadership, consensus decision making and collective behaviour in humans," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, pp. 781--789, 2009.
- [5] J. Berg and C. H. Karud, Swarm intelligence in bio-inspired robotics. *Master's thesis, Norwegian University of Science and Technology*, Norway, 2011.
- [6] I. M. Rekleitis, G. Dudek, and E. E. Miliotis, "Multi-robot exploration of an unknown environment, efficiently reducing the odometry error," in *International Joint Conference on Artificial Intelligence*, 1997, pp. 1340-1345.
- [7] I. A. D. G. A. M. E. Rekleitis, "Multi-robot collaboration for robust exploration," *Annals of Mathematics and Artificial Intelligence*, vol. 31, no. Springer, pp. 7--40, 2001.
- [8] R. C. Arkin and J. Diaz, "Line-of-sight constrained exploration for reactive multiagent robotic teams," *7th*

- International Workshop on Advanced Motion Control*, 2002, pp. 455-461.
- [9] T. A. A. R. C. Balch, "Communication in reactive multiagent robotic systems," *Autonomous Robots*, vol. 1, no. Springer, pp. 27-52, 1994.
- [10] M. Powers and T. Balch, "Value-based communication preservation for mobile robots," *In Distributed Autonomous Robotic Systems 6*, Springer Japan, 2007, pp. 327-336.
- [11] Nguyen, Hoa G., et al., Maintaining communications link for a robot operating in a hazardous environment. *Space And Naval Warfare Systems Command San Diego CA*, 2004.
- [12] Howard, Andrew, Lynne E. Parker, and Gaurav S. Sukhatme, "Experiments with a large heterogeneous mobile robot team: Exploration, mapping, deployment and detection," *The International Journal of Robotics Research*, vol. 25, no. 5-6, 431-447, 2006.
- [13] B. Yamauchi, "Frontier-based exploration using multiple robots," *in Proceedings of the second international conference on Autonomous agents, ACM*, 1998, pp. 47-53.
- [14] R. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. Younes, "Coordination for multi-robot exploration and mapping," *in Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence*, 2000.
- [15] C. Stachniss, M. Moos, and W. Burgard, "Efficient exploration of unknown indoor environments using a team of mobile robots," *Annals of Mathematics and Artificial Intelligence*, vol. 52, pp. 205-227, 2008.
- [16] K. M. Wurm, C. Stachniss, and W. Burgard, "Coordinated multi-robot exploration using a segmentation of the environment," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008.
- [17] S. Koenig, B. Szymanski, and Y. Liu, "Efficient and inefficient ant coverage methods," *Annals of Mathematics and Artificial Intelligence*, vol. 31, pp. 41-76, 2001.
- [18] I. Rekleitis, G. Dudek, and E. Milios, "Multi-robot collaboration for robust exploration," *Annals of Mathematics and Artificial Intelligence*, vol. 31, no. Springer, pp. 7-40, 2001.
- [19] M. Schwager, B. Julian, and D. Rus, "Optimal coverage form multiple hovering robots with downward-facing cameras," *In International conference on Robotics and Automation*, Japan, Kobe, 2009.
- [20] Dyer, John RG, et al., "Leadership, consensus decision making and collective behaviour in humans," *Philosophical Transactions of the Royal Society B: Biological Sciences*, pp. 781-789, 2009.

Comparative Study of Testing Tools Blazemeter and Apache Jmeter

Pirah Memon¹, Tahseen Hafiz², Sania Bhatti², Saman Shahid Qureshi¹

Abstract:

Automated Testing plays a vital role in the entire development of software. Due to growing requirements of an automated testing diverse range of testing tools are available. From literature, it is observed that a number of automated testing tools are studied and compared. However, this is the first time that Apache JMeter and BlazeMeter are compared. The objective of this paper is to compare load testing tools: Apache JMeter and BlazeMeter based on the criteria such as performance, latency, size, error percent, duration count, number of hits and response time. This paper focuses on the analysis of the performance and functionality that minimizes the software cost and resources. After performing experiments, it is proved that the performance of BlazeMeter is better than Apache JMeter with respect to all parameters.

Keywords: *BlazeMeter; Apache JMeter; Automated software testing.*

1. Introduction

Software testing is used to find out the errors in the software product. Software testing identifies the product completeness, correctness and also used to improve the product quality. Testing does not guarantee the error-free software. Rather, testing helps to debug the error within the software. There are various approaches for testing, which depend upon the software requirements, category of software and available resources. In simple word, testing is to “verify the product and evaluate it”. Where the term verifies is something that the tester wants to match the requirements with the actual product and the response of product with the behavior in action to the analysis of the tester. Although most of the intellectual properties are identical to the inspections of the requirements the word testing is concerned with dynamic analysis of the product. Testing helps us improve the quality of the software. The quality of the software can be improved by involving the non-functional attributes

according to the standard of ISO-9126. Testing is about verifying and validating software if it is working as it is intended to design. It involves testing of a product using static and dynamic methodologies because of human mistakes, manual designs. Hence the quality of the software can be achieved by performing the quality assurance activities. It is usual for the developer to spend 40% of the software cost on testing. For example bank transaction monitor, control can cost 3 to 5 times as much as all other activities are combined due to the antagonistic nature of testing the developer does not consider the notations in its development of software.

2. Related Work

There has been a significant work on the comparative analysis of HP LoadRunner and Apache JMeter in literature but limited work on BlazeMeter. The limitation of the paper proposed by V.Chandel [1] is that they only described the Apache JMeter HP LoadRunner but does not depict the real time results. The

¹ Institute of information and communication Technology Mehran University of Engineering and Technology Jamshoro, Pakistan

² Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan
Corresponding Email: pirahmemon01@gmail.com

automated web services testing tools presented in study [2] are very interesting as it describes the Apache JMeter, SoapUI and Storm in detail but they do not justify that among them which tool is better. In paper [3] M.S Sharmila. et al. Discussed the Apache JMeter in a well mannered way but the factors for comparison with other tools are not focused. In study [4] K. Tirghoda illustrated the Apache tool in detail but any script for web services which are being tested using the Apache tool is not generated. Sadiq et al. [5] uses response time, throughput, latency, scalability and resource utilization for Apache JMeter but does not delineate the security issues related to Apache JMeter. B. Patel, et al. [6] compared two performance testing tools i.e LoadRunner and JMeter. They compared the parameters such as load generating capacity, installation, download proficiency result reporting, cost, technicality of software and reliability. The comparative analysis is done between HP LoadRunner and Apache JMeter by R. B. Khan [7] but author only targets the websites LOAN calculator and BMI calculator because of not enough traffic on these websites. Authors in [8] put light on two load testing tools: Sikuli and Commercial Tool for acceptance testing. They compared on static properties and industrial traffic management system but there is no statistical difference between these tools; both performed the same automated testing. In study [9] the empirical analysis of web service testing tools is performed with the technical features and the comparison is completed on the basis of performance only. In a recent study [10] the comparative analysis is done among the Selenium, Soap UI, HP QTP/UFT and Test Complete on the basis of different features. Authors use the 3-point scale, i.e. good, average and bad in comparison. The results are presented in the form of graphs based on the calculated values for selected tools and soap is considered as the best tool among them.

3. Apache JMeter

Apache JMeter [11] is an Apache open source software, a pure Java application, designed to perform the functional behavior and performance testing specially on web applications that further expand into the other applications on both static and dynamic resources (web services SAOP/RESET) web dynamic languages PHP, Java and Asp.net files but JMeter does not execute the Java script found within the html pages nor it does render the html pages as browser does. It can also be used to graphically analyze the performance or to test the server\script\object behavior under heavy concurrent load.

Figure 1 shows that as the request is sent by the user is directly acknowledged by the server, the server responses to the user's request, then JMeter collects the data and manipulates the statistical information of further task will be completed and the results will be displayed.

Figure 2 defines the test plan for starting testing of facebook. Test plan requires four elements http default request, http, cookie manager, a listener and graph result.

Figure 3 describes the number of users along with the test start and end time with a loop count at each thread while the users are generated.

In figure 4, the black labels show the data of the facebook, the blue line defines the slight change in average number of users, pink line defines the no change occur in the median while testing. However, red line shows that deviation changing occur constantly, whereas the green line output increases as the number of users increases. On the y-axis the maximum time is 63068 milliseconds is the time required for completing the facebook uniform resource locator test. Here the graph result depicts the visual model of facebook samples that can be read and written from a file.

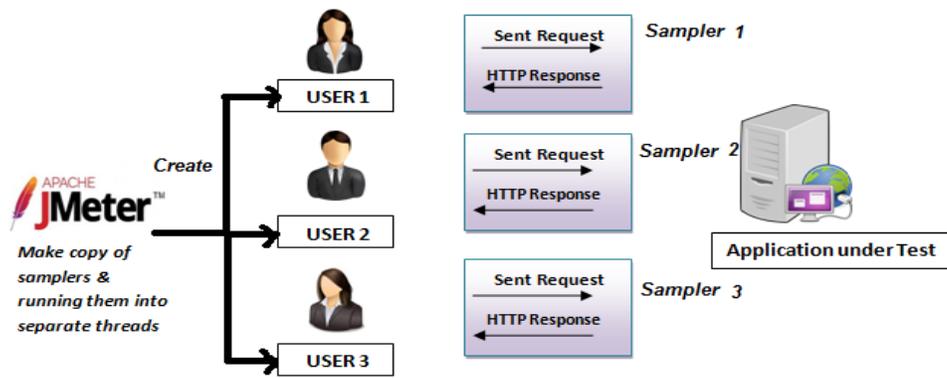


Fig. 1. Working mechanism of Apache JMeter [3].

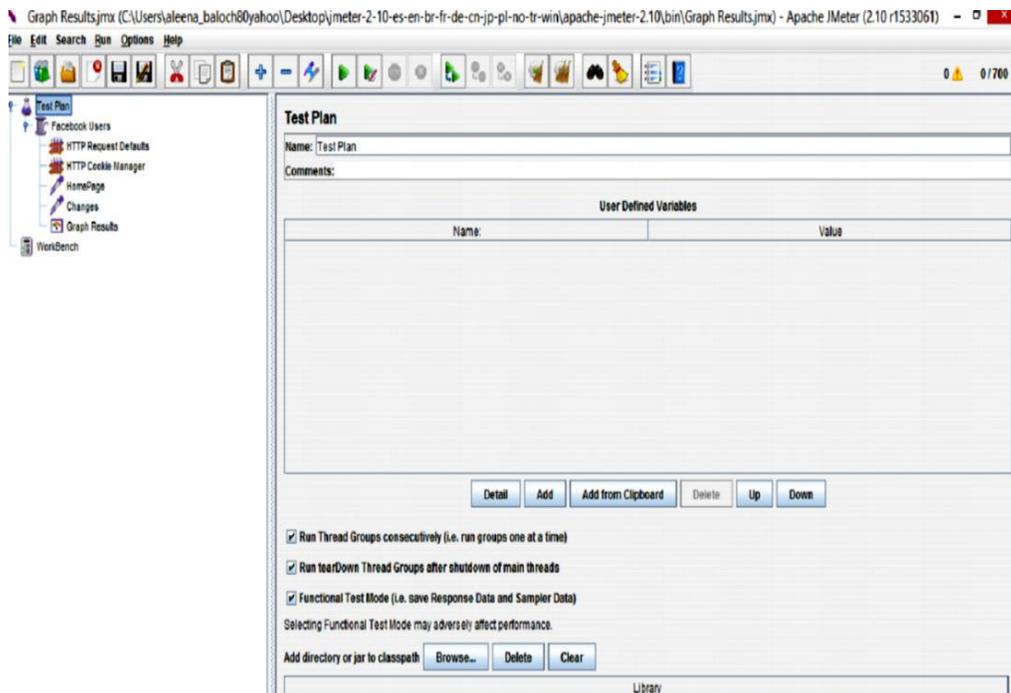


Fig. 2. Facebook url test with Apache JMeter.

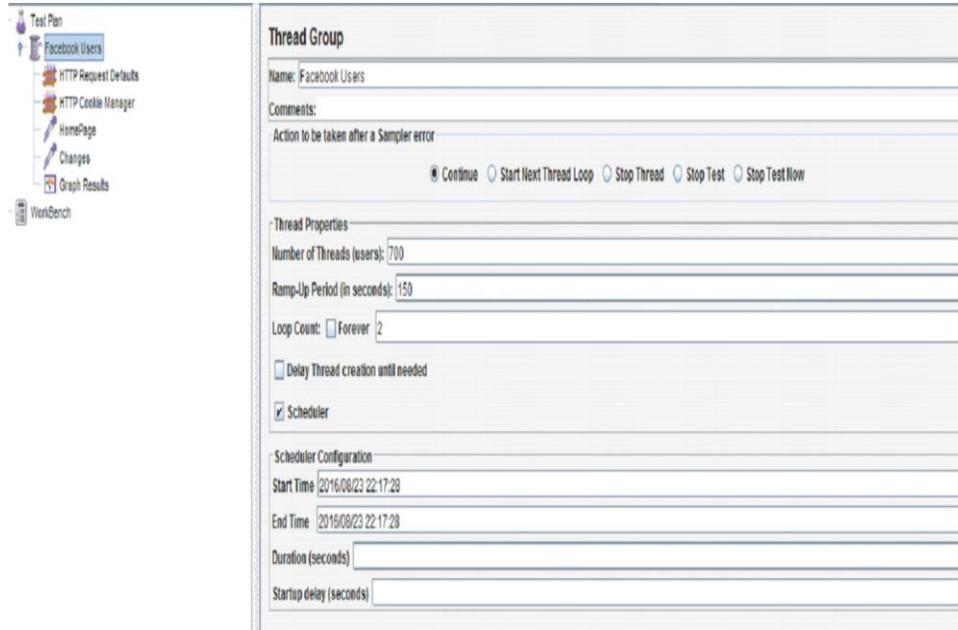


Fig. 3. Thread group of facebook url test with Apache JMeter.



Fig. 4. Graphical depiction of facebook url test.

4. Blazemeter Testing Tool

BLAZEMETER [12] is an enterprise tool fully compatible with Apache JMeter. It provides the developer's simple integrated tool into their native development environment. Blaze Meter is used to test the mobile, web applications; web site, web services, and database testing that can simulate thousands of users. The objective of this paper is to conduct a comparative analysis of social website, facebook to improve the performance of the social websites by making the system more reliable at a less response time.

5. Results And Discussion

The goal of this work is to improve the performance of social website "facebook" by performing load testing with BlazeMeter instead of Apache JMeter. Because Apache JMeter fails when the scalability of the product is increased and the behavior is inflexible that means modifications cannot be applied after performing testing even if they are needed. The performance of the testing tools is judged on the basis of the following factors.

Flexibility: It is the non- functional quality attributes of software engineering. Flexibility refers to how easily the changes can be accommodated in the system.

Scalability: Scalability refers to the how easily the system can be expanded by increasing the number of users.

Performance: Performance is the ability of the system to perform the task.

Load Controller: The device used to regulate the amount of power that a load can consume. It can be used by third party energy or utility to reduce the customer energy demands at the certain time.

Reliability: The ability of the system to perform failure free operation for a specified time in a specified environment.

Aggregate Reports: It allows reviewing an overview of administrative information for various settings and status.

Latency Time: It is defined as the amount of time a message takes to reverse a system or to reach a designation.

Table 1. Comparison of testing tools based on various factors.

	<i>Factors</i>	<i>BlazeMeter</i>	<i>Apache JMeter</i>
1	Flexibility	Yes	Yes
2	Scalability	Yes	Yes
3	Performance	Yes	Yes
4	Load controller	Yes	No
5	Reliability	Yes	Yes
6	Aggregate reports	Yes	No
7	Latency time	No	Yes

Table 02, as described in Section IV, defines that at each point the virtual number of users is generated with minimum response time at an average bandwidth of 20KB with zero percent error and also shows that the test pass successfully.

TABLE 2. Result of facebook url test with 0% error and 50 virtual users.

Max. Users	Avg. Throughput	Error	Avg. Response Time	90% Response Time	Avg. Bandwidth
50VU	0.78 Hits/s	0 %	174.5 (MS)	211 (MS)	20.15 KB

Figure 5 shows the 0% error means the test pass successfully with fifty users and maximum number of hits which are 0.86 per milliseconds.



Fig. 5. Result of facebook url test with BlazeMeter for number of hits.avg.

TABLE 3. Samples of Facebook URL test with response time.

Element Label	Samples	Avg. Response Time (Ms)	Avg. Hits/s	99%Line (Ms)	Min. Response Time (ms)	Avg. Bandwidth Bytes/s	Error Rate
Facebook	19612	187.3	16.34	624	155	2931	0%



Fig. 6. Result of facebook BlazeMeter for latency url test with BlazeMeter for latency.

Figure 6 depicts that at each point the test is carried by the virtual users with maximum throughput and average response time of 157 milliseconds.

In Table 3 column 2 describes the number of samples with value 19612, column 3 shows the average response time to value of 187.3 milliseconds, column 4 depicts the average number of hits, column 5 defines the time, in milliseconds, column 6 illustrates the minimum response time, column 7 the average bandwidth and column 8 represents the error 0% means test pass successfully.

6. Conclusion

Currently, software testing has become the necessity for the organizations. Because it saves both time and money. Apache JMeter and BlazeMeter are very efficient for encountering the performance testing of software. From experiments it is apparent that BlazeMeter tool is more efficient as compared to Apache JMeter. It has a simple, clean User interface that shows what's going on without confusion And too much effort and it offers straightforwardness with its Uniqueness. Moreover, it is free of cost and possesses effective Portability with 100% Java purity. Both of them are open Source projects and have merits, but neither is ideal. Because the experiments are performed with a limited number of users and more experiments are required to be performed with increased number of users.

REFERENCES

- [1] V. Chandel et al., "Comparative Study of Testing Tools: Apache JMeter and Load Runner," *International Journal of Computing and Corporate Research*, 2013.
- [2] G. Murawski, et al., "Evaluation of load testing tools," 2014.
- [3] M. S. Sharmila and E. Ramadevi, "Analysis of performance testing of web application," *International Journal of Advanced Research in Computer and Communication Engineering*, 2014.
- [4] K. Tirghoda, "Web Services Performance Testing Using Open Source Apache Jmeter," *International Journal of Scientific & Engineering Research*, vol. 3, 2012.
- [5] M. Sadiq, et al., "A Survey of Most Common Referred Automated Performance Testing Tools," *ARN Journal of Science and Technology*, vol. 5, pp. 525-536, 2015.
- [6] B. Patel, et al., "A Review Paper on Comparison of SQL Performance Analyzer Tools: Apache JMeter and HP LoadRunner," 2014.
- [7] R. B. Khan, "Comparative Study of Performance Testing Tools: Apache JMeter and HP LoadRunner," Ed. 2016.
- [8] E. Borjesson and R. Feldt, "Automated system testing using visual GUI testing tools: A comparative study in the industry," in *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, 2012, pp. 350-359.
- [9] S. Sharma and A. K. Sharma, "Empirical Analysis of Web Service Testing Tools".
- [10] M. Imran, et al., "A Comparative Study of QTP and Load Runner, Automated Testing Tools and their Contributions to Software Project Scenario," 2016.
- [11] A. S. Foundation. 2017-1-19. Apache JMeter. Available: <http://jmeter.apache.org/>.
- [12] A. Girmonsky. (2016, 01-12-2016). BlazeMeter. Available: <http://blazemeter.com/>.

Internet of Things (IoTs) for Disaster Management

Syeda Ambreen Zahra¹, Iqra Shafique¹, Tuba Farid¹

Abstract:

It is critical that rescuers can track the caught victims and perform composed help actions quickly. The ordinary media transmission framework (e.g. a landline or cell system) might be either mostly or totally harmed by a catastrophe occurrence. Internet of Things (IoTs) is an encouraging innovation that can be utilized to take care of a portion of the issues said above. To date, the use of IoT in disaster administration is as yet an unexplored issue. The target of this paper is to concentrate the IoT-based proposition for disaster administration structure.

Keywords: *IoT; Disaster; Crowdsourcing; Cloud computing; WSN.*

1. Introduction

Natural and man-made disasters, for example, quakes, surges, plane accidents, elevated structure falls, or major atomic office breakdowns, represent an ever-display test to open crisis administrations. Disaster organization has been attracting a lot of thought by many research gatherings, including Computer Science, Environmental Sciences, Health Sciences and Business. The makers' gathering starts from a product building establishment, and particularly from the zone of data organization and examination. Keeping in mind the end goal to adapt to such disasters in a quick and very organized way, the ideal arrangement of data concerning the circumstance is a basic pre-imperative [3].

Police, fire offices, general wellbeing, common guard and different associations need to respond effectively and exclusively, as well as in an organized way. This outcomes in the requirement for both intra and inters association coordination at a few order levels [1]. Since management requires current data and such data must be reported upstream and downstream inside and between associations progressively, the need emerges for an integrated communication and data framework for disaster administration that gives proficient, solid and secure trade and

preparing of important data. Regardless of the starting, crisis conditions are consistently joined by instability of how the disaster will develop, a sharp pace of response operations, and the probability of honest to goodness loss of human lives and property if not responded to really.

Other grouping plans exist, however whatever the cause, certain elements are fascinating for administration of all disasters Prevention, Advance alerted, Early acknowledgment, Analysis of the issue, and examination of degree, Notification of the overall public and fitting masters, Mobilization of a response, Containment of damage, Relief and helpful watch over those impacted because of natural change, among various causes, disastrous occasions have extended out and out consistently and that is not simply costing us to the extent assets/system hurt more over in dynamic adversities of human lives [2]. While we cannot stop the occasion of basic disasters, with the help of present day advancement one can extra people's lives more enough. Exchanges systems in the midst of a damaging occasion can be the complexity among life and downfall for those in the impacted areas [5].

¹ CS & IT, University of Lahore, Gujrat, Pakistan

Corresponding Email: Msituol@gmail.com

2. Idea of IoTs (Internet of things)

There are various implications of the Internet of Things in the investigation and critical present day gatherings. The definitions may climb from the word 'Web 'and provoke an 'Internet organized 'vision, or 'things 'and incite a 'things arranged 'vision. Putting the world 'Internet' and 'Things' together semantically suggests a general arrangement of interconnected differences strangely addressable, in perspective of standard correspondence traditions. The term Internet of Things (IoT) has been around for a long time [3].

In this circumstance, it is gaining ground with the improvement of exploiting edge remote advancement. The basic idea of this thought is the proximity of a variety of articles – for instance, RFID, NFC, sensors, actuators, mobile phones, et cetera which, through uncommon tending to arrangements, can work together with each other. Presently a day's assorted advancements of IoT, for example, RFID (Radio Frequency Identification), Near Field Communication (NFC), Machine-to-Machine Communication (M2M) and Vehicular-to-Vehicular correspondence (V2V) are there in the business parts which are used to execute the front line thought of IoT [4].

3. Purpose Behind Picking IoT

Over the span of regular framework breakdown, D2D correspondence to be begun and confined an uncommonly selected framework where a segment of the devices will go about as a hand-off or portal administrator. This hand-off administrator will interface the impacted domain with rest of the world at whatever point they get any live advances, for instance, Wi-Fi, Satellite or working standard cell coordinate. Using IoT we will make a work arrangement of different devices so if one device can't confer then another device will be in used so there is no delay and objectivity among correspondence frameworks, IoT will be an assorted framework and distinctive sorts of devices will be related there [7].

Among them, general devices including equipment and devices for different IoT application spaces, for instance, mechanical machines, home electrical devices, sharp vehicles and pushed cells and so forth. These general devices may viably be embedded with high taking care of and computational chipsets, and hence may talk with various frameworks paying little character to the individual developments used.

4. Some Portion of IoT in Disaster Management

It was energized by the way that the Internet has transformed into our exchanges spine for the web, and things rise toward phone calls. If there should be an occurrence of a disaster, power can go out, servers can go down and systems can wind up obviously over-weight, all of which can impact Internet-based exchanges. Regardless of having some redundancy and support structures set up, it is probably not going to expect that we could ever make the Internet truly impenetrable to any disaster [8]. When you consider the Internet of Things (IoT), the quickly growing number of devices in our lives that can connect with the Internet and to each other, you doubtlessly consider the ways it can make your life less requesting [6].

For example, the IoT starting now empowers us to do things like control the indoor controllers in our homes using an application on our phones. Regardless, adjacent to the solaces it can offer, the (IoT) Internet of Things moreover can serve an essential, possibly lifesaving, and part in the event of calamity, typical or something else. Today, immense scale catastrophe slant and response requires organizations which rely on the information that is being secured in these barely detectable IoT frameworks. These frameworks, however starting at now set up, are ceaselessly being revived, modernized, and pervaded with the latest advances which will over the long haul transform into the qualification between sensible disappointment

and groundbreaking pulverization for countless [9].

The purpose behind this paper is to give a comprehension into the current IoT based work and highlight a possible response for post – calamity response organization. The rule nature of this paper is the accentuation that IoT can be a promising one, represent an ever – display test to open crisis administration.

Whatever is left of the paper is sorted out as takes after. Area we expose the idea of IoT. Then we talk about purposed methodology and related work. At long last, we display the conclusions and distinguish open research difficulties to build up an IoT based disaster flexible correspondence arrange.

5.Literature Review

Several solutions have been proposed by researchers to adequately maintain communication after disaster.

1. Jeva et al ., proposed a system called **“DBAPRS (Disaster Behavior Analysis and Probabilistic reasoning System)”** for notifying future disaster alleviation and management in the city of Japan [13]. It can also figureout at what percent people will migrate to numerous cities of Japan in case of disaster. This system can also examine people’s evacuation behavior at the time of Great East Japan Earthquake.

2. According to Nan Jing, **“Context – Aware Disaster Response System”** can classify and inspect the setting data of mobile application users [34]. The major drawback of this framework was that Sky guard did not consider security measures when to accumulate and explore the assessment of mobile users.

3. Asli Soyler proposed that the structure and conduct of a disaster management System can be acquired in a single demonstrating environment by utilizing both the model based framework and its modeling language [35].

4. Sarmad Sadik et al., proposed an architecture called **“Policy based Migration of Mobile agents”** that has been utilized for managing the exertion execution and the moving pattern of Mobile agents [36]. This model can also moderate the movement of mobile agents to specific areas and regulate the execution of certain actions on source and target machines.

5. According to Li Zbigangt et al., **“Urban disaster management information system”** can enhance the reaction speed and exactness of Government Emergency management [37]. Through this system, analyzing disaster information, managing safety measure and administration of crisis become feasible.

6. Hassan et al., presents a Novel model that comprises of Mobile Cloud (MC) called **“D2D based Mobile Cloud”**. On the basis of Residual energy and signal to noise interface (SNR), user Equipment’s (UE’s) challenge the cluster heads (CH’s) [32].

Numerous legitimate areas (Clusters) have different cluster heads (CH’s). Contrast to traditional mobile based communication, this model shows an increase of 25% in bandwidth and data transfer proficiency [14].

7. Ahmed et al., proposed a software **“Arc GIS Simulation tool”** for anticipating the upcoming disaster and also analyze pre and post disaster flood risk analysis and an Ad hoc Wireless sensor network (WSN) architecture [16]-[31]. This software is very useful in emergency situation and also help rescuers to take preventive actions for saving the life of victims in case of critical conditions. The proposed architecture comprises the following three subtypes

- WSN Area
- GIS based Emergency Response DB Server
- Remote Sensing and Satellite based infrastructure

For predicting and analyzing Flood analysis, this research demands integration

with Wireless Sensor Network (WSN). Using proposed software, they had also performed simulations for estimating flood in different regions of Sindh. The GIS enabled Map for flood forecasting gives a broad insight about the areas which are likely to be affected from heavy rainfall during recent two years [15].

8. Anthone et al., proposed an alternative solution of sending SMS in case of natural disaster. This system was called “**Alternative Emergency SMS network**”. This system is very handy for users as it provides an alternative means of communication. The proposed system is based on sending SMS directly to the Short Message Service Center (SMSC) utilizing the SMS interchange conventions over Wireless Mesh Sensor Network (WMSN) [16]- [33].

9. Ashish Rauniyar et al., proposed a model in IOT called “**CDMFC (Crowdsourcing Disaster Management Fog Computing)**”. The model further classified into four layers namely [24]:

- **Sensing Layer**

The Purpose of this layer is to sense both natural and artificial disasters like; flood, fire, earthquake and many IOT based applications with the help of different sensors, mobile phone, Laptops and tablets, etc. The purpose of this layer is to only produce sensing information rather than the kind of event.

- **Crowdsourcing Layer**

The main purpose of this layer is to crowdsource the data that is being sensed from the above layer (Sensing layer). Then this data is transferred to the cloud for detailed investigation in which different strategies like data mining are being tested to make this data understandable.

- **CDMFC Layer**

This layer accommodate filtering techniques that depends upon emergency and keywords (relevant to disaster). Development

of these keywords is being possible by the IoT applications, through humans using different mobile phones, tablets and sensors dispose in disaster affected district or area.

With the help of crowdsourcing and data offloading mechanism, disaster relevant IoT information explored in CDMFC layer in an effective manner. Facebook and twitter generated data gives information of areas and time stamps progressively. The exact location and time of disaster in a shorter time span can be identified from the information generated through Facebook and twitter [10].

This layer also comprises urgent contact numbers. These numbers are directly approachable by rescuers, who can arrange safeguard in case of disaster. This will be helpful in sense to make vital move as per crowdsourced basic calamity related IOT information.

- **Cloud Computing Layer**

This layer stores and inspects all the basic and non- basic information that is being generated from Crowdsourced layer.

10. Devasena et al., proposed two types of sensors for measuring the disaster [25]:

- **Homogenous WSN**

In this network, sensor nodes have identical attributes. They measure the same sort of parameters such as temperature. Examples of the clusters that are being intended for homogeneous WSN are Hybrid Energy Efficient Distributed clustering (HEED), Power efficient gathering in sensor information system (PEGASIS) and low Energy Adaptive clustering Hierarchy (LEACH).

- **Heterogeneous WSN**

In this network, sensor nodes have distinctive attributes because they need to quantify diverse parameters. SN are independent to take decisions to fulfill sensing task, building topology for network and routing policies. In this way, it ends up

noticeably essential to outline energy efficient algorithm for upgrading robustness against node failures and extending lifetime of WSN [11].

Heterogeneous WSN can be useful in disasterprone regions because it can interpret and analyze more than one parameters. It can also be fundamental to figure out the upcoming disaster and relevant actions could be taken on the basis of information produced by these clustering protocols.

It additionally includes geographic graphs in which areas are digitally entered by address, interpret with calculation that delivers a likelihood surface indicating the possibility where crime incidents are high [12].

6. Comparison & Result

It is a fact that all nations bear the effect of disaster directly or indirectly. Consequently, different techniques used in different ways are to be compared. In order to achieve this objective, different methods and parameters are used in this paper. The table 1 compares all techniques, limitations, event phase and type of disaster of different methods.

7. Proposed Methodology

In this paper, we have figureout that the systematic simulation and forecast of all types of disaster are conceivable. The principle challenge is that we have to propose such a mechanism that will sort out and examine the context information of mobile application users and utilizing that information in a manner to customize context aware and targeted instruction to mobile user.

In our review paper, we recommend “CDFMC” for disaster management. This model has an advantage of Fog computing platform. Basic crowdsourced IOT disaster relevant information is examined progressively through this platform. This model also asserts data offloading mechanism, when a direct link to Fog computing is not available. By employing block chain technology, offload mechanism send disaster

related IOT data to the CDFMC/fog layer [24]-[27]. We can also figure out disasters in actual time and manage plans for rescue operations with the help of this model.

In addition, information relevant to disaster interpret on Fog and remaining information inspect on cloud. Hence this model can converse the bandwidth. In comparison to cloud computing model which are much more likely to be targeted by attackers to employ IoT data, this model would be beneficial to run the IoT data securely within Fog where the user can introduce their own incompetent Security Algorithms.

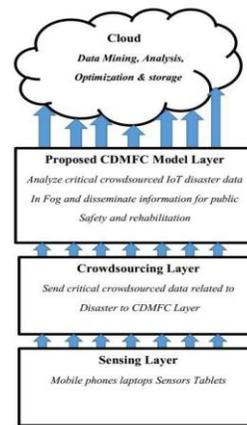


Fig.1. Hierarchical Layered Structure of CDMFC model for disaster Management [24]

8. Conclusion

Disasters are of increasing frequency and extremity in the current world. For the past years wireless technology has seen a massive progress in communications. To provide the command, control, and communications abilities needed in emergency situations, public safety and emergency management organizations increasingly rely on wireless technology [38].

In this paper, disaster and disaster management are defined by using IoT. Based on the results, we can say that natural disasters can be supervised if they are correctly managed, and acceptable infrastructures can control the disaster before it becomes ruin and

preparedness before disasters can remarkably reduce losses. Finally, criteria for disaster management is given that could be helpful for further disaster management planning and disaster studies [39].

TBALE 1. Comparison and Results.

TECHNIQUE	LIMITATIONS	EVENT PHASE	TYPES OF DISASTER
DBARAS (Disaster Behavior Analysis and Probabilistic Reasoning System)	Difficulty in analyzing moving patterns	Ultimate disaster relief & management	Earth quakes, Tsunami
Context Aware System Disaster Response System	Security Issues when storing and investigating mobile user data	Disaster response phase	All type of Disaster
Policy based Migration of Mobile Agents	Collaboration issue among mobile agents	Disaster response phase	Earthquakes
Urban Disaster Management Information System	Complex research required	Disaster response phase	Fire, Cyclone & Flood etc.
Arc GIS Simulation Tool	Complex research required	Future prediction and management phase	Flood
Crowdsourcing Disaster Management Fog Computing	Latency & Security	Disaster response phase	
Alternative Emergency SMS Network	Reliability Security	Disaster response phase	Natural Disaster
D2D based Mobile Cloud Architecture	Coverage Restrictions	Disaster response phase	All type of Disaster
Homogenous & Heterogeneous			Tsunami, storm, volcano & Erath quake

ACKNOWLEDGMENT

We authors acknowledge with thanks the assistance rendered by Mr. Aziz Bhatti Sheikh, University of Lahore, Gujrat Campus

for providing crucial insight during the course of the research work which greatly improved the manuscript.

REFERENCES

- [1] N. Ahmad and M. Hussain, "Flood Prediction and Disaster Risk Analysis using GIS based Wireless Sensor Networks," *International Journal of Basic and Applied Scientific Research*, pp. 2090-4304, 2013.
- [2] S. Ray and J. Gutierrez, "A Study of IoT-based Post-Disaster Management," *ICOIN*, vol. 7, no. 19, pp. 406-409, 2017.
- [3] V. Hristidis and Y. Deng, "Survey of data management and analysis in disaster situations," *The Journal of Systems and Software*, vol. 7, no. 19, pp. 1701-1707, 2010.
- [4] O. Said and M. Masud, "Towards Internet of Things: Survey and Future Vision," *International Journal of Computer Networks (IJCN)*, vol. 5, no. 1, pp. 3-9, 2013.
- [5] C. Aggarwal and N. Ashish, "The Internet of Things: A Survey from the Data-Centric Perspective," vol. 1, pp. 2-10, 2013.
- [6] S. Salman and F. Solehria, "A Proposed Least Cost Framework of Irrigation Control System Based on Sensor Network for Efficient Water Management in Pakistan," *International Journal of Basic & Applied Sciences IJBAS-IJENS*, vol. 11, no. 2, 2011.
- [7] Er. Yadav and Er. Ankur "The Internet of Things: Impact and Applications in the High-Tech Industry," vol. 1, pp. 3-9, 2015-2016.
- [8] [Online]. Available: http://us.corwin.com/sites/default/files/ubinary/6244_Chapter_4_Boba_Final_PDF_3.pdf. [Accessed: 23- Nov- 2016].
- [9] A. Balte and A. Kashid, "Security Issues in Internet of Things (IoT): A Survey," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol.5, pp. 5-9, 2015.
- [10] S. Fang et al., "An Integrated System for Regional Environmental Monitoring Management Based on Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1596-1605, 2014.
- [11] M. Ancona et al., "An Internet of Things Vision of the Flood Monitoring Problem," *The Fifth International Conference on Ambient Computing, Applications, Services and Technologies*, 2015, vol. 3, pp. 26 – 29.
- [12] M. D. Kamruzzaman et al., "A study of IoT- Based Post - Disaster Management," *Research Gate*, pp. 406 – 410, 2017.
- [13] M. A. Khoshkholghi et al., "Disaster Recovery in Cloud Computing: A Survey," *Computer and Information Science*, vol. 7, no. 4, pp. 39 – 54, 2014.
- [14] I. Priyadarshinee et al., "Flood Prediction and Prevention through Wireless Sensor Network (WSN): A Survey," *International Journal of Computing Applications*, vol. 113, no. 9, pp. 30 – 36, 2015.
- [15] L. Atzori et al., "The Internet of Things: A Survey," *Computer Networks*, no. 54, pp. 2787 – 2805, 2010.
- [16] N. Altay et al., "OR/MS Research in Disaster Operation Management," *European Journal of Operational Research*, no. 175, pp. 475 – 493, 2006.
- [17] E R. Pooja and ER. Ankur, "A Survey of Growth and Opportunity of Internet of Things (IoT) in Global Scenario," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 12, pp. 20664 – 20671, 2016.
- [18] D. S. Rao et al., "A Survey of Emergency Communication Network Architectures," *International Journal of u – and e – Service, Science and Technology*, vol. 8, no. 4, pp. 61– 68, 2015.
- [19] M. Yazdanipour, "Survey on Disaster Management in Developing Countries and Developed Countries," *UACEE International Journal of Computer Science and its Applications*, vol. 2, no. 3, pp. 51 – 55, 2013.

- [20] A. Saoji and P. Lambhate, "Survey Paper on Event Detection Techniques in Wireless Sensor Network for Disaster Recovery," *International Journal of Engineering Research & Technology (IJERT)*, vol. 2, no. 12, pp. 120 – 124, 2013.
- [21] V. R. Jeva and J. Puthiyidam, "A Survey on Disaster Management System," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 2, no. 11, pp. 355 – 360, 2014.
- [22] A. Rauniyar et al., "Crowdsourced – Based Disaster Management using Fog Computing in Internet of Things Paradigm," *IEEE 2nd International Conference on Collaboration and Internet Computing*, 2016, pp. 490 – 494.
- [23] A. Devasena and B. Sowmya, "Wireless Sensor Network in Disaster Management," *Indian Journal of Science and Technology*, vol. 8, no. 14, pp. 1 – 6, 2015.
- [24] C. Perera et al., "Sensing as a Service Model for Smart Cities Supported by Internet of Things," *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 1, pp. 81–93, 2014.
- [25] F. Jalali et al., "Fog Computing May Help to Save Energy in Cloud Computing," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, pp. 1728-1739, 2016.
- [26] Z. Xu et al., "Crowdsourcing Based Description of Urban Emergency Events using Social Media Big Data," 2016.
- [27] M. Aazam and E.-N. Huh, "Fog Computing: The Cloud-IoT/IoE Middleware Paradigm," *IEEE Potentials*, vol. 35, no. 3, pp. 40–44, 2016.
- [28] T.E. Drabek, "Theories Relevant to Emergency Management versus a Theory of Emergency Management," *Journal of Emergency Management*, vol. 3, no. 4, pp. 49-4, 2005.
- [29] N. Ahmad et al., "Flood Prediction and Disaster Risk Analysis using GIS based Wireless Sensor Networks: A Review," *Journal of Basic and Applied Scientific Research*, ISSN 2090-4304, 2013.
- [30] S.T.Hassan et al., "Mobile Cloud based Architecture for Device – to – Device (D2D) Communication Underlying Cellular Network," *Research Gates*, 2015.
- [31] V. Anthone and S. Oishi , "A wireless mesh sensor network framework for river flood detection which can be used as an emergency communications network in case of disaster," *11th International Conference on Hydro informatics HIC*, 2014.
- [32] N. Jing and Y. Li, "A Context-aware Disaster Response System Using Mobile Software Technologies and Collaborative Filtering Approach," *IEEE 18th International Conference on Computer Supported Cooperative Work in Design*.
- [33] A. Soyler, "A Model-Based Systems Engineering Approach to Capturing Disaster Management Systems,"
- [34] S. Sadik et al., "Policy Based Migration of Mobile Agents in Disaster Management Systems," *2nd International Conference on Emerging Technologies*, 2006, Peshawar, Pakistan, pp. 13-14.
- [35] Li. zbigangt et al., "Research of GIS-based Urban Disaster Emergency Management Information System," *International Conference on Computer and Communication Technologies in Agriculture Engineering*, 2010.
- [36] D. Srinivasa rao et al, "A Survey of Emergency Communication Network Architectures," *International Journal of u- and e- Service, Science and Technology*, vol.8, no.4 , pp. 61-68, 2015.
- [37] E. J. Catlos et al, "Nepal at risk: Interdisciplinary lessons learned from the april 2015 nepal (gorkha) earthquake and future concerns," *GSA Today*, vol. 26, no. 6, 2016.
- [38] L. B. Brengarth and E. Mujkic, "Web 2.0: How social media applications leverage

- nonprofit responses during a wildfire crisis,” *Computers in Human Behavior*, vol. 54, pp. 589–596, 2016.
- [39] A. R. Lee, A. Rauniyar, and S. Y. Shin, “Implementation of escarpment alarm system using terrestrial reference system,” *International Journal of Future Computer and Communication*, vol. 4, no. 1, p. 72, 2015.

0.7" ↑ Margin Top

Paper Formatting Guidelines

Title: 14^{pts} Bold

Author Name: 11^{pts} Bold

Affiliation: 11^{pts} Italic

Abstract: Single Paragraph (Min: 150 – Max: 250 words)

Paper Length: Formatted as guided (Min: 4,000 – Max: 8,000 words)

Font: Times New Roman 10^{pts}

Font Color: Black

Line Spacing: 1.0 (single line space throughout the paper)

Indentation: Justify

References & In-text Citations: Follow the IEEE & Use EndNote X7 for In-text citations and Bibliography.

Headings: 12^{pts} Bold

1. 12^{pts} Bold

1.1. 11^{pts} Bold

1.1.1. 11^{pts} Bold Italic

← Margin Left
0.8" →

← Margin Right
0.5" →

6.5" x 10"

Page Size: 6.5" (Width) x 10" (Height)

Submission: Formatted paper as guided can be submitted through our online submission system at <http://sjcms.iba-suk.edu.pk>

1.9" ↑ Margin Bottom
↓

Publisher: **Sukkur IBA Journal of Computing and Mathematical Sciences (SJCMS)**

Office of Research, Innovation & Commercialization – ORIC

Sukkur IBA University - Airport Road Sukkur-65200, Sindh Pakistan

Tel: (09271) 5644233 Fax: (092 71) 5804425 Email: sjcms@iba-suk.edu.pk URL: sjcms.iba-suk.edu.pk

Sukkur IBA **Journal** of Computing and Mathematical Sciences



SUKKUR IBA UNIVERSITY
Merit - Quality - Excellence

SUKKUR IBA UNIVERSITY
Airport Road, Sukkur -65200
Sindh, Pakistan
Tel: +92-71-5644233
Fax: +9271-5804419
Email: sjcms@iba-suk.edu.pk
URL: sjcms.iba-suk.edu.pk