TEQIP–II Sponsored

International Conference on Communication Computing & Systems

(ICCCS-2014)

8th-9th August 2014

Editors
Dr. Satvir Singh
Dr. Vishal Sharma
Dr. Krishan Saluja
Dr. Monika Sachdeva

Shaheed Bhagat Singh State Technical Campus
Moga Road (NH–95), Ferozepur-152004, Punjab, India

International Conference on
COMMUNICATION, COMPUTING & SYSTEMS
(ICCCS–2014)

Editors
Dr. Satvir Singh
Dr. Vishal Sharma
Dr. Krishan Saluja
Dr. Monika Sachdeva

Shaheed Bhagat Singh State Technical Campus
Moga Road (NH–95), Ferozepur-152004, Punjab, India

EXCEL INDIA PUBLISHERS
NEW DELHI

First Impression: 2014

International Conference on Communication, Computing & Systems (ICCCS–2014)

ISBN: 978-93-83842-68-1

DISCLAIMER

The authors are solely responsible for the contents of the papers compiled in this volume. The publishers or editors do not take any responsibility for the same in any manner. Errors, if any, are purely unintentional and readers are requested to communicate such errors to the editors or publishers to avoid discrepancies in future.

Published by
EXCEL INDIA PUBLISHERS
91 A, Ground Floor
Pratik Market, Munirka, New Delhi–110067
Tel: +91-11-2671 1755/ 2755/ 3755/ 5755 ● Fax: +91-11-2671 6755
E-mail: publishing@groupexcelindia.com
Web: www.groupexcelindia.com

Typeset by
Excel Publishing Services, New Delhi–110067
E-mail: prepress@groupexcelindia.com

Printed by
Excel Printing Universe, New Delhi–110067
E-mail: printing@groupexcelindia.com

Foreword

Shaheed Bhagat Singh State Technical Campus (Formerly, known as SBS College of Engineering & Technology), Ferozepur was established in 1995 by the Government of Punjab. Institute is serving our nation by producing quality engineers and researchers. Well qualified, self-motivated and experienced faculty is always busy in their research, teaching and overall development of students.

SBS State Technical Campus, Ferozepur is covered under World Bank assisted Technical Education Quality Improvement Programme, Phase-II (TEQIP-II) by the MHRD and Government of Punjab. In contrast, three department, viz. Electronics & Communication Engineering, Computer Science & Engineering and Electrical Engineering have collectively taken an initiative to conduct this first ever International Conference on Communication, Computing and Systems (ICCCS–2014) to encourage young researchers.

During the project, a web based conference management system is developed. Every received paper has been assigned a paper identification number, authors registered themselves, and final camera ready papers are collected through same software. Emails are also sent from time to time so as to share information among large group of authors.

More that hundred research papers were received from various domains of research. Originally, we had a plan of publishing only fifty papers. Therefore, review process was very rigorous involving three blind reviews and an online plagiarism check. However, during the review process, our reviewer could not limit them to only fifty papers.

Numerous paid journals on various topics have been established in the market as business. Many of them do not bother about the quality and contents of the research paper, except fee. Our Conference Team does not want to promote this business, therefore, not associated any paid journal with this conference. However, all research papers published in this Conference Proceeding will also be made available online at www.sbsstc.ac.in/icccs2014. Any researcher can download any paper, free of cost, to study and refer in their present research. This will also increase citations of published papers in conference proceeding.

Conference Team has done its job with best possible efforts. Suggestions from the readers and authors are always welcome. We will definitely take care in future.

With Best Wishes

Conference Team

Acknowledgments

Shaheed Bhagat Singh State Technical Campus has taken a historical leap to conduct a first ever *International Conference on Communication, Computing and Systems–2014* (ICCCS–2014). This is culmination of joint efforts of three departments, viz. (1) Electronics & Communication, (2) Computer Sc. & Engineering and (3) Electrical Engineering.

SBS State Technical Campus, Ferozepur is covered under World Bank assisted Technical Education Quality Improvement Programme, Phase-II (TEQIP-II) by the MHRD and Government of Punjab. This conference is a sincere effort of all our faculty members, students and contributing researchers toward improving our educational and research standards in contrast to the TEQIP-II Project.

Our sincere thanks are due to National Project Implementation Unit (NPIU) for inclusion of our institute in World Bank assisted project of TEQIP. Our gratitude to Sh. B. Purushartha IAS, Director and Sh. Mohanbir Singh, Additional Director, DTE&IT and State Project Facilitation Unit (SPFU) Punjab who gave us necessary guidelines from time to time. We also want to acknowledge Dr. Anand K. Tyagi, TEQIP-II Coordinator and his TEQIP Team who supported us financially from TEQIP funds in organizing this event.

We could never reach at this end without the support of Dr. Tejinder Singh Sidhu, Director, Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab. He motivated and pushed us with unconditional support to provide all the facilities from the institute.

Conference Team wants, on the behalf of all our colleagues, to thank Sh. Dinesh Lakra, Chairman, Board of Governors, SBS State Technical Campus, Ferozepur for his valuable suggestions and sustained support to organize such activities those can earn name and fame for our institute.

We want to acknowledge our contributing authors and co-authors. Their unexpected huge response in the form of research papers increased our excitement. From total received research papers, due to organizational and publication constraints, about 40% papers were required to be dropped. This is why only the best research papers were selected, after three blind reviews and aplagiarism check, for publication in the Conference Proceedings. It was a difficult job for Review Team. We feel sorry for non-acceptance of some of papers although they were good.

Finally, we must thank GOD for giving us the environment to organize this event, people to help, opportunities to encash and potential to succeed.

Conference Team

Message

I am very happy to be a part of this International event being organized at Shaheed Bhagat Singh State Technical Campus, Ferozepur. I humbly accept the honor of being the Chief Guest in the *International Conference on Communication, Computing Systems (ICCCS–2014).*

The students, researchers, scientists and giants in industry will share their knowledge and findings with their counterparts and discuss on this excellent platform. The Conference will be a step forward to achieve self-reliance and be a country of innovators.

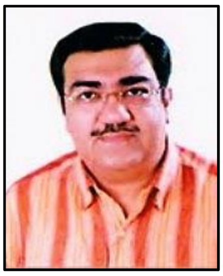
I hope that the SBS state campus will set the ball rolling and will organize more such conferences from time to time and it will be a periodic feature of this great SBS state campus. The Researchers, Developers and people involved in Industries will share their experiences on the platform being provided henceforth here.

This type of activities will encourage the students to become JOB providers rather than JOB seekers. I wish the Conference a great success, and some useful results so attained here will be conveyed to quarters concerned for implementation.

With best wishes

Professor (Dr.) Tara Singh Kamal
FIE, FIETE, Life Sr. MIEEE (USA)
Vice President (2001–02), Chairman (1999–2001)
IE (India) Punjab & Chandigarh State Center
President (2004–05)
Engineering Sciences Section,
Indian Science Congress Association

Message



We all associates of the SBS State Technical Campus are one family. I being the head of this family congratulate all on this auspicious occasion of first ever *International Conference on Communication, Computing and Systems (ICCCS–2014)* being organized in our campus. This is the beginning of new era in the progressive path of development of our campus and it is a matter of proud for us that after attaining excellence at undergraduate level we are ready to excel in the field of Higher Education and Research.

As this collaborative project of World Bank and Government of India the Technical Education Quality Improvement Programme, Phase-II (TEQIP-II) is being implemented in the campus and its outcomes are getting evident. The volume and quality of research papers received indicate this region has the ability to become the hot bed of research activates.

I expect now onwards, ICCCS will be a periodic event to provide an International Platform for Researchers and Students to present their work.

With Best Wishes

Dinesh Lakra
Chairman BOG
SBSSTC Ferozepur

Message

Shaheed Bhagat Singh State Technical Campus is organizing first *ever International Conference on Communication, Computing and Systems (ICCCS–2014)* and, on this occasion, it gives me immense pleasure to send my greetings to the participants of ICCCS and good wishes to the organizers for the success of the conference.

By organizing this kind of events it seems that this institute is ready to play its role in the field of higher education and research as well. This initiative is in line with the policy of Government of India to spread the higher education at mass level. The research activities like, this conference, are very much essential for upliftment of higher education level in the country. These activities will create a positive environment for research potential of the region to develop.

I am sure that this conference would through up useful ideas on a topic which is important for our present as well as future generations.

I extend my best wishes to the organizers and the participants on this occasion.

Dr. T.S. Sidhu
Director
SBSSTC Ferozepur

Message

It is a matter of great pleasure and proud to me that ICCCS–2014 Conference is being organized at this institute and the proceedings of the same are under published in the form of a book. I offer my congratulations to the participants, speakers, organizing and TEQIP team on this occasion. The aim of ICCCS–2014 is to bring together leading academic scientists, researchers and research scholars to exchange and share their experiences and research results about all aspects of Electrical, Computer, Electronics and Communication Engineering. It also provides the premier interdisciplinary and multidisciplinary forum for researchers, practitioners and educators to present and discuss the most recent innovations, trends, and concerns, practical challenges encountered and the solutions adopted in the field of Electrical, Computer, Electronics and Communication Engineering.

I am highly thankful to Hon'ble Chairman BOG Sh. Dinesh Lakraji and Worthy Campus Director Prof. T.S. Sidhu ji who had always been very sensitive to the issues concerning the upliftment of this institution, for their valuable cooperation and patronage in organizing this conference. If we are in position to offer this wonderful event today to the academic fraternity, it is simply because of the excellent men and women who have staffed this institution and also who have been trained here all these years. What SBSSTC has done over last more than eighteen years now–"in caritate et justitia", that is, in love and fairness towards all-is to create a culture of excellence and hardwork. SBSSTC has been what we call as "magis", a term meaning a constant and relentless struggle to go beyond, to push further and further the contours of our own limitations. Eighteen years is a crucial period as it signifies the "adulthood" whereas in the life of an institution, it is the time to prove its worth and I feel glad that SBSSTC has made its mark on the state's psyche. I am sure that this conference will positively prove to be a step forward in this direction.

Professor Anand K. Tyagi
TEQIP Coordinator &
Associate Director (Engineering Wing)

Message

I am feeling delighted for being the Convener of this first ever *International Conference on Communication, Computing and Systems (ICCCS–2014)* being organized at Shaheed Bhagat Singh State Technical Campus, Ferozepur. Organizing an international conference for me is like a dream come true. Receiving research papers, managing their reviews, keeping authors informed and publishing accepted research paper in form of Conference Proceeding, etc. are some of very critical activities that I learnt during the process.

This conference is actually resultant of joint efforts of all the faculty members of three department (1) Electronics & Communication Engineering, (2) Computer Science & Engineering and (3) Electrical Engineering. Everyone in the organizing team was propelled by unconditional support from Director and Board of Governors and their contribution made this event a great success.

After attaining excellence at undergraduate level SBSSTC Ferozepur is now ready to push research and innovation standards in the region. This conference is just a beginning and I am sure in future such events will be regular activities of our institute.

Best Wishes

Dr. Satvir Singh
Convener

Core Organizing Team

Conveners

Dr. Satvir Singh
Convener

Dr. Vishal Sharma
Co-convener

Coordinators

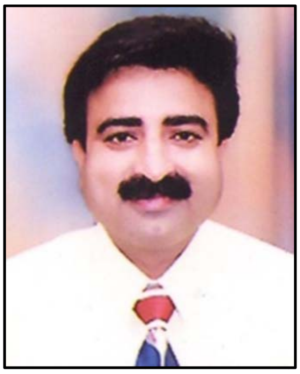
Dr. Krishan Saluja
Coordinator

Dr. Monika Sachdeva
Coordinator

Dr. Kultar Deep Singh
Coordinator

Co-coordinators

Mr. Japinder Singh
Co-coordinator



Mr. Sanjeev Dewra
Co-coordinator

Editorial Board

Organizing Committees

Publication Committee

1. Dr. Satvir Singh, Chief Editor
2. Dr. Vishal Sharma, Editor
3. Dr. Krishan Saluja, Co-Editor
4. Dr. Monika Sachdeva, Co-Editor

Technical Support Committee

1. Mr. Anil Bansal
2. Mr. Parminder Pal Singh
3. Mr. Amardeep Chopra

Boarding & Lodging Committee

1. Dr. V. S. Bhullar
2. Dr. Anila Gupta
3. Ms. Rajni
4. Mr. Japinder Singh
5. Mr. Gurnam Singh
6. Mr. Vishal Arora

Discipline Committee

1. Dr. Lalit Sharma
2. Dr. V.S. Bhullar
3. Dr. Tejeet Singh
4. Mr. Sanjeev Dewra
5. Mr. Gurjeevan Singh
6. Mr. Baldev Singh Mann

Refreshment Committee

1. Mr. Gurmeet Singh
2. Mr. Ravi Kant Grover
3. Mr. Amarjeet Singh
4. Mr. Jaswant Singh

Transportation Committee

1. Mr. Sukhwant Singh
2. Mr. Satinder Sharma
3. Mr. Naresh Kumar

Venue Arrangement Committee

1. Dr. Kiranjeet Kaur
2. Dr. Anila Gupta
3. Dr. Neel Kanth Grover
4. Mr. Gulshan Ahuja
5. Mr. N.S. Bajwa

Mementos and Certificate Committee

1. Mr. Vishal Arora
2. Mr. Ram Pal
3. Mr. Parvesh Puri

Reception Committee

1. Dr. Krishan Saluja
2. Dr. Sangeeta Sharma
3. Ms. Daljeet Kaur
4. Ms. Anupam Mittal
5. Ms. Jawinder Kaur
6. Mr. Anil Bansal

Registration Committee

1. Dr. Krishan Saluja
2. Mr. Sarabjeet Singh
3. Ms. Navdeep Kaur
4. Mr. Chakshu Goel
5. Mr. Amardeep Chopra
6. Mr. Amit Grover

Finance Committee

1. Mr. Inder Pal Singh
2. Mr. Japinder Singh
3. Mr. J.K. Aggarwal
4. Mr. Sat Pal

Medical Committee

1. Dr. Ajay Kumar
2. Mr. Kamal Bhatti

Contents

Comparative Analysis of EDDEEC & Fuzzy Cost Based EDDEEC Protocol for WSNs

Baljinder Kaur[1] and Parveen Kakkar[2]
[1,2]Department of Computer Science & Engineering,
DAV Institution of Engineering & Technology, Jalandhar, India
E-mail: [1]bkkhinda@gmail.com, [2]parveen.daviet@gmail.com

***Abstract*—This research paper has focused on evaluating the performance of the heterogeneous WSNs using fuzzy based cluster head selection and the tradition EDDEEC protocols. The EDDEEC has used different probability function for choosing the best cluster head by using the remaining energy and average energy of the network. But EDDEEC has neglected the use of number of neighbours of sensor nodes during cluster head selection. Whereas fuzzy based EDDEEC heterogeneous protocol is based on the concept of thee fuzzy cost. The fuzzy cost is dynamic in nature and evaluated on the basis of the remaining energy and the node centrality. MATLAB simulation tool is considered in this paper. The comparative analysis has shown that the fuzzy cost based EDDEEC outperforms over the existing EDDEEC protocol.**

Keywords: Network Lifetime, Stable Period, Clustering, DEEC, Fuzzy Cost

I. Introduction

Wireless sensor networks (WSNs) [1] are among the widely used types of ad-hoc wireless networks. Main objective of WSNs is to classify, collect, and development of the information within a monitoring area. In 1980's DARPA i.e., Defence Advanced Research Projects Agency started a program named Distributed Sensor Networks (DSN) from which further WSN was formed. WSN consists of more than hundreds of little sensor nodes which have restricted power, memory and computational capability. These node route data and throw it base station called as sink. For communication of data among nodes and sink many routing technology are used firstly, such as direct communication and multi-hop data communication [3]. But due to restricted battery life of nodes this techniques were not so useful because early loss of some node in both techniques were be unsuccessful to acquire in the network appropriateness periods. The purpose of the WSN involves many fields such as the armed field; reforest fire finding, earthquake detection, air pollution structure monitor and other intense environments. The sensor nodes in WSN have restricted power, memory and computational capability. A sensor node makes use of its communicating mechanism in order to transmit the data, over a wireless channel, to a base station (sink). WSNs accept energy-constrained battery-powered devices. The sensor nodes are abounding by non-rechargeable batteries mount on sensors, therefore minimize energy utilization in order to expand the existence of network is an important issue in WSNs [8]. Since the major portion of energy utilization in sensor nodes is due to communication, variety of a capable algorithm considerably reduces the communication energy. By clustering of sensor nodes into some groups called clusters, sensor nodes of each cluster send their data to definite sensor nodes in the cluster called Cluster Heads (CHs). Then, CH nodes spread gathered information to the Base Station (BS). A sensor network design is affected by many factors [9], which include scalability, fault tolerance, sensor network topology, production costs, transmission media, operating environment, hardware constraints, and power consumption. These factors are important because they provide guideline to an algorithm or design various routing protocol [10] to improve the network lifetime of WSNs.

II. Various Clustering Techniques

A. Leach

Low Energy Adaptive Clustering Hierarchy i.e., LEACH [4] is the first hierarchical cluster-based routing set of rules for wireless sensor network. In LEACH, the nodes classify themselves into local clusters. A dedicated node preferred as cluster-head is dependable for designing and employing a TDMA (Time Division Multiple Access) plan and aggregating the data coming from different nodes and sending it to the BS [6]. The process of LEACH is divided into round. In this protocol each round has two phases: Set-up Phase and Steady-state Phase.

B. Teen

The main features of Threshold sensitive energy efficient sensor network [5] protocol are that the sensor nodes have to send out to their CH to consume fewer energy, extra data processing is done only by CH to reduce energy consumption, CHs deployed at higher levels of hierarchy transmit the data which use more energy. To remove this issue, all nodes are given the opportunity to be CH for a time period T (cluster Period). In TEEN, nodes sense the network all the time and data broadcasting is done only when there is an extensive change in the sensed data. HT is the absolute value of an attribute to trigger on its transmitter and report to its respective CH. HT allows nodes to broadcast data, if the data occurs in the range of interest. Therefore, a considerable reduction of the transmission delay occurs. Moreover, ST is the small change in the value of the sensed attribute. Next transmission takes place when there is a small change in the sensed attribute once it arrives at the HT. So, it further decreases the number of transmissions [7].

C. DDEEC

Developed Distributed Energy Efficient Clustering (DDEEC) [13] protocol uses same way for judgment of usual energy in the network and CH collection algorithm based on remaining energy as deployed in DEEC [11]. Difference between DDEEC and DEEC is centred in phrase that defines probability for a normal and an advanced node to become a CH.

It is found that nodes with more remaining energy at round r are more likely to become CH, in these way nodes having advanced energy principles or advanced nodes will become CH more number of times as compared to the nodes with low energy. After some time in a sensor network there comes a point where there are advanced nodes having same remaining energy level like normal nodes energy level are present in the network, still after this point also DEEC continues to penalize the advanced nodes. So this is not the best way to allocate the energy as advanced nodes are constantly becoming CH and due to this they die early than the normal nodes. SEP [16] is used only for two level heterogeneous sensor network. To avoid this uneven case, DDEEC makes some changes to keep away advanced nodes from being punished over and again. It introduced Threshold residual energy [13] in which all type of nodes has same chance to become CHs for current round.

$$\mathrm{T}h_{REV} = E_0\left(1 + \frac{aE_{disNN}}{E_{disNN - E_{disAN}}}\right) \tag{1}$$

D. EDEEC

Enhanced Distributed Energy Efficient Clustering (EDEEC) [14] used perception of three stage heterogeneous network. It contains three types of nodes normal, advanced and super nodes based on original energy. p_i is the possibility used for CH collection and p_{opt} is indication for p_i.

$$p_i = \frac{p_{optE_{i(r)}}}{1+m(a+m_{0b)E(r)}} \text{ if } \mathrm{s_i} \text{ is the normal node} \tag{2}$$

$$p_i = \frac{p_{opt(1+a)E_{i(r)}}}{1+m(a+m_{0b)E(r)}} \text{ if } \mathrm{s_i} \text{ is the advanced node} \tag{3}$$

$$pi = \frac{p_{opt(1+b)E_{i(r)}}}{1+m(a+m_{0b)E(r)}} \text{ if } \mathrm{s_i} \text{ is the super node} \tag{4}$$

E. TDEEC

Threshold Distributed Energy Efficient Clustering (TDEEC) [17] uses similar method for CH choice and usual energy evaluation as proposed in DEEC. At each about nodes has been decided whether it can become a CH or not by choosing an arbitrary number between 0 and 1. If number is less than threshold then nodes choose to become a CH for the given round. In TDEEC, threshold value is used to and based upon that rate a node decides whether to become a CH or not by introducing remaining energy and average energy of that round with respect to best possible no of nodes.

F. EDDEEC

Enhanced Developed Distributed Energy Efficient Clustering (EDDEEC) [2] method is used for heterogeneous WSNs. It is three level heterogeneous WSNs. It uses same scheme for CH choice based on initial, remaining energy level of the nodes, radio dissipation and average energy of network as in DEEC. At beginning of the round, each node makes a decision whether to become a CH or not for current round based on Threshold. heterogeneous Wireless sensor network have more than two types of nodes so in EDDEEC three level heterogeneity are used which contain normal, advance and super nodes and uses same probabilities of three types of nodes as described in EDEEC [14]. In EDEEC after some rounds, some super and advance nodes have same remaining energy level as normal nodes due to continually CH selection. Therefore it continues to penalize advance and super sensor nodes for CH choice. Same issue with DEEC, it also continues to penalize just advance nodes and DDEEC is limited only for two-level heterogeneous networks. To eliminate this unbalanced problem in three-level heterogeneous WSNs EDDEEC changes in function which illustrated in EDEEC for calculating probabilities of normal, advance and super nodes. These modifications are based on absolute remaining energy level $\mathrm{T_{absolute}}$, that is the value in which advance and super sensor nodes have similar energy level as in case of normal nodes. Using $\mathrm{T_{absolute}}$ all kinds of nodes has identical probability for CH selection.

$$\frac{\mathrm{P_{opt}E_{\,i}(r)}}{\mathrm{(1+m(a+m_{\,o}b))\bar{E}(r)}}$$
$$\text{for normal nodes if } \mathrm{E_{\,i}(r) > T_{absolute}}, \tag{5}$$

$$\frac{\mathrm{P_{\,opt}(1+a)E_{\,i}(r)}}{\mathrm{(1+m(a+m_{\,o}b))\bar{E}(r)}}$$
$$\text{for advance node if } \mathrm{E_{\,i}(r) > T_{absolute}}, \tag{6}$$

$$\frac{\mathrm{P_{\,opt}(1+b)E_{\,i}(r)}}{\mathrm{(1+m(a+m_{\,o}b))\bar{E}(r)}}$$
$$\text{for super nodes if } \mathrm{E_{\,i}(r) > T_{absolute}}, \tag{7}$$

$$c\,\frac{\mathrm{P_{\,opt}(1+b)E_{\,i}(r)}}{\mathrm{(1+m(a+m_{\,o}b))\bar{E}(r)}} \text{ Otherwise} \tag{8}$$

Here $\bar{\mathrm{E}}(r)$ is average energy at round r of the network, $\mathrm{E_{\,i}(r)}$ is residual energy at round r, m is fraction between node heterogeneity, P_{opt} is probability of optimum number of cluster head, a, b is boost a power for advance and super nodes.

III. Fuzzy Based EDDEEC

Step 1: Initialize the WSNs with required parameters like nodes position, sink position, initial energy of each kind of nodes etc.

Step 2: for every node i repeat the following steps until all nodes become dead.

Step 3: Select cluster head using following equations i.e. normal (eq. 9), advance (eq. 10), super nodes (eq. 11) and for all types of nodes having same remaining energy (eq. 12).

$$\frac{P_{opt}E_i(r)}{(1+m(a+m_o b))\bar{E}(r)} * \text{Fuzzy\_cost for normal nodes if } E_i(r) > T_{absolute} \quad (9)$$

$$\frac{P_{opt}(1+a)E_i(r)}{(1+m(a+m_o b))\bar{E}(r)} * \text{Fuzzy\_cost for advance node if } E_i(r) > T_{absolute} \quad (10)$$

$$\frac{P_{opt}(1+b)(r)}{(1+m(a+m_o b))\bar{E}(r)} * \text{Fuzzy\_cost for super nodes if } E_i(r) > T_{absolute} \quad (11)$$

$$c\frac{P_{opt}(1+b)E_i(r)}{(1+m(a+m_o b))\bar{E}(r)} * \text{Fuzzy\_cost for nor, adv, sup nodes if } E_i(r) \leq T_{absolute} \quad (12)$$

Where Fuzzy cost will be evaluated using the Algorithm 1.

Step 4: Evaluate the energy dissipation and update the remaining energies it. Where distance will be evaluated using eq. 5 and updating of energy will be based upon the eq. 6 and eq. 7.

$$d_{toCH} = \frac{M}{\sqrt{2\pi k}},\ d_{toBS} = 0.765\frac{M}{2} \quad (13)$$

$$E_{Tx}(l,d) = l E_{elec} + l\varepsilon_{fs} d^2,\ d < d_0 \quad (14)$$

$$E_{Tx}(l,d) = l E_{elec} + l\varepsilon_{mp} d^4,\ d \geq d_0 \quad (15)$$

Algorithm 1: Fuzzy Cost Calculation

I. Initial Round:

1. Sink selects cluster head according to the weighted probability function (Padv for advance nodes, Psup for super nodes, Pnrm for normal nodes) and broadcast the CH_message.
2. Cluster formation will be done and data transform will be take place.
3. Each sensor node computes the remaining energy and sensor node centrality and sends the values to sink through Cluster Head.
4. End

II. General Rounds:

1. Fuzzy cost will be calculated by Sink using remaining energy and sensor node centrality.
2. Sink chooses the Cluster Head based on the value of fuzzy cost and broadcast the CH_message.
3. Cluster formation will be done and data transform will be take place.
4. Each sensor node computes the sensor node centrality and remaining energy and sends the values to sink through Cluster Head.
5. End

IV. Experimental Setup

This section contains the experimental setup which has been used in this research paper. Table 1 has shown various constants and variables required to simulate this work. These parameters are standard values used as benchmark for WSNs.

Table 1 Experimental Setup

Parameter	Value
Area (x,y)	100,100
Base Station (x,y)	50,50
Nodes (n)	100
Probability (p)	0.1
Initial Energy (Eo)	0.1
Transmiter_Energy	50 nJ/bit
Receiver_Energy	50 nJ/bit
Free space (Amplifier)	10 nj/bit/m^2
Multipath (Amplifier)	0.0013 pJ/bit/m^4
A (Energy Factor Between Normal and Super Nodes)	3
B (Energy Factor Between Normal and Advance Nodes)	2
Maximum Lifetime	5000
Message Size	4000 bits
M (Fraction of Advanced Nodes)	0.3
X (Fraction of Super Nodes)	0.3
Effective Data Aggregation	5 nJ/bit/signal

V. Experimental Results

On applying fuzzy cost functions, following results will be achieved using MATLAB tool.

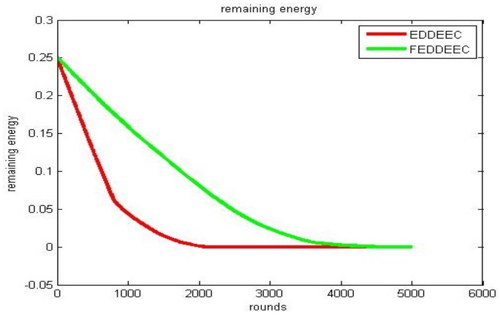


Fig. 1 Remaining Energy

Figure 1 is showing the comparative analysis of remaining energy. X-axis is representing the energy in joules. It has been clearly shown that the remaining energy in cased of FEDDEEC is quite more than that of the EDDEEC. Thus FEDDEEC outperforms over the EDDEEC with respect to remaining energy.

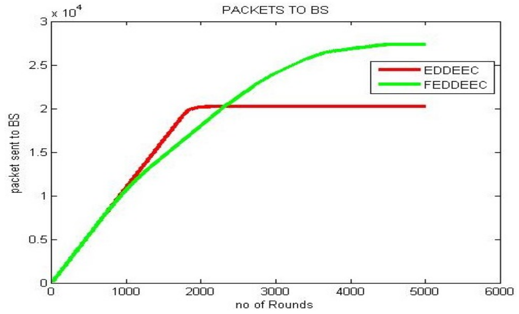


Fig. 2 Total Number of Packets Sent to Base Station

Figure 2 is showing the comparison of FEDDEEC and the EDDEEC with respect to total number of packet sent to base station. X-axis is representing packet sent to base station. Y-axis is representing number of rounds. It has been clearly shown that the overall packers sent to base station in case of FEDDEEC are quite more than that of the EDDEEC. Thus FEDDEEC outperforms over the EDDEEC with respect to packets sent to base station.

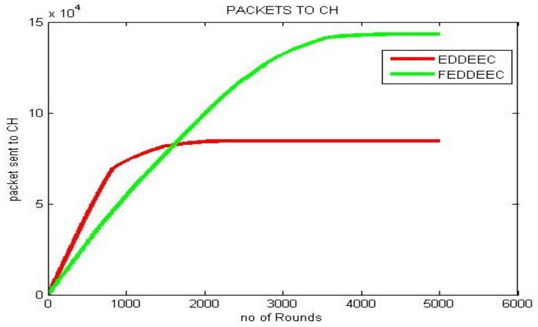


Fig. 3 Total Number of Packets Sent to Cluster Head

Figure 3 is showing total number of packet sent to cluster head. It has been evidently shown that the overall packers sent to cluster head in instance of FEDDEEC are fairly supplementary than that of the EDDEEC. Thus FEEDDEEC overtakes over the EDDEEC with respect to packets sent to cluster head.

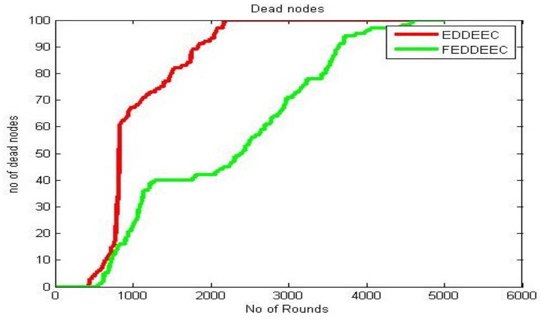


Fig. 4 Total Number of Dead Nodes

Figure 4 is showing total number of dead nodes during the network lifetime. It is showing all nodes die at 2172 and 4601 round respectively. Thus FEDDEEC overtakes over the EDDEEC with respect to dead nodes. This figure has shown that Fuzzy based EDDEEC increased the network lifetime.

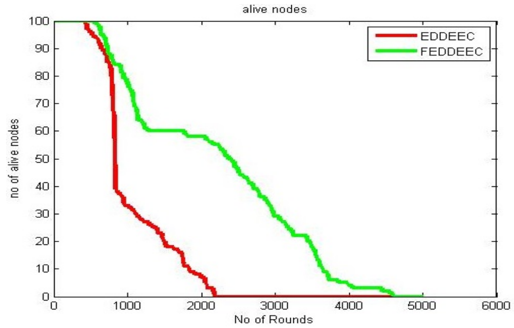


Fig. 5 Stability Period

Figure 5 is showing the stability period of the nodes. It is showing that the first node for EDDEEC and FEDDEEC dies at 423 and 517 round, respectively. This figure has shown that Fuzzy based EDDEEC increased the stability period.

TABLE 2 COMPARATIVE ANALYSIS

Protocols	First Node Dead	Last Node Dead
EDDEEC	423	2172
FEDDEEC	517	4601

Table 2 has shown the comparison between EDDEEC and FEDDEEC with respect to first node dead and last node dead time.

VI. CONCLUSION

This paper has focused on the performance analysis of EDDEEC and fuzzy cost based EDDEEC. The FEDDEEC is based on fuzzy cost and has the ability to overcome the limitation of the EDDEEC by optimized dividing the sensor field among consistent number of clusters. Due to the limitation of the real time environment this work has done simulation in the well-known MATLAB tool. The comparative analysis has shown that the due to the fuzzy based optimization in the FEDDEEC it significantly improve the results than that of existing EDDEEC in terms of packet sent to base station, network lifetime and stability period. In near future we will justify the proposed algorithm further by using the mobile sink and also by placing the sink statically in and outside the sensor field.

REFERENCES

[1] S.B. ALLA *et al.*, "Balanced and Centralized Distributed Energy Efficient Clustering for heterogeneous wireless sensor networks", 3rd International Conference on Next Generation Networks and Services, 2011, pp. 39–44.

[2] N. Javaid *et al.*, "EDDEEC: Enhanced Developed Distributed Energy-Efficient Clustering for Heterogeneous Wireless Sensor Networks.", International Workshop on Body Area Sensor Networks, 2013, pp. 914–919.

[3] M. Alshowkan *et al*., "LS-LEACH: A New Secure and Energy Efficient Routing Protocol for Wireless Sensor Networks", 17th International Symposium on Distributed Simulation and Real Time Applications (DS-RT), 2013, pp. 215–220.

[4] W.R. Heinzelman *et al*., "Energy-efficient communication protocol for wireless microsensor networks", in Proc. 33rd Hawaii IEEE International Conference on System Sciences, 2000, pp.1–10.

[5] A. Manjeshwar and D.P. Agrawal., " TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks", 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, 2001, pp. 1–7.

[6] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," IEEE Transaction on Wireless Comm., Vol. 1, No. 4, pp. 660–670, 2002.

[7] Manjeshwar, Arati, and Dharma P. Agrawal. "APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks.", Proceedings of the 16th International Parallel and Distributed Processing Symposium (IPDPS), 2002, pp. 1–8.

[8] S. Lindsey and C.S. Raghavendra, "PEGASIS: Power-efficient gathering in sensor information systems", In Aerospace conference proceedings, 2002, pp. 1125–1130.

[9] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks", IEEE Transaction on Mobile Computing, Vol. 3, No. 4, pp. 366–379, 2004.

[10] G. Smaragdakis, *et al*., "SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor network", In Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA), 2004, pp. 1–10.

[11] L. Qing *et al*., "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks", Computer communications, Vol. 29, No. 12, pp. 2230–2237, 2006.

[12] N. Israr and I. Awan, "Multihop clustering algorithm for load balancing in wireless sensor networks", International Journal of Simulation Systems, Science and Technology, Vol. 8, No. 1, pp. 13–25, 2007.

[13] B. Elbhiri *et al*., "Developed Distributed Energy-Efficient Clustering (DDEEC) for heterogeneous wireless sensor networks", 5th International Symposium on Communications and Mobile Network, 2010, pp. 1–4.

[14] P. Saini and A.K. Sharma, "E-DEEC-Enhanced Distributed Energy Efficient Clustering Scheme for heterogeneous WSN", 1st International Conference on Parallel, Distributed and Grid Computing (PDGC), 2010, pp. 205–210.

[15] A. Kashaf *et al*., "TSEP: Threshold-sensitive Stable Election Protocol for WSNs", 10th International Conference on Frontiers of Information Technology (FIT), 2012, pp. 164–168.

[16] Y. Miao *et al*., "Performance Study of Routing Mechanisms in Heterogeneous WSNs", International Conference on Computer Science & Service System (CSSS), 2012, pp. 971–974.

[17] T.N. Qureshi *et al*., "On Performance Evaluation of Variants of DEEC in WSNs",7th International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA), 2012, pp. 162–169.

Technologies to Handle Big Data: A Survey

Sabia[1] and Love Arora[2]
[1,2]Department of Computer Science & Engineering,
Guru Nanak Dev University, Regional Campus, Jalandhar, India
E-mail: [1]sabiajal@gmail.com, [2]aroralove7@gmail.com

***Abstract*—Big data came into existence when the traditional relational database systems were not able to handle the unstructured data (weblogs, videos, photos, social updates, human behaviour) generated today by organisation, social media, or from any other data generating source. Data that is so large in volume, so diverse in variety or moving with such velocity is called Big data. Analyzing Big Data is a challenging task as it involves large distributed file systems which should be fault tolerant, flexible and scalable. The technologies used by big data application to handle the massive data are Hadoop, Map Reduce, Apache Hive, No SQL and HPCC. These technologies handle massive amount of data in MB, PB, YB, ZB, KB and TB. In this research paper various technologies for handling big data along with the advantages and disadvantages of each technology for catering the problems in hand to deal the massive data has discussed.**

Keywords: Big Data, Hadoop, Map Reduce, Apache Hive, No SQL

I. INTRODUCTION

With the growth of technological development and services, the large amount of data is formed that can be structured and unstructured from the different sources in different domains. Massive data of such sort is very difficult to process that contains the information of the records of million people that includes everyday massive amount of data from social sites, cell phones GPS signals, videos etc. Big data is a largest buzz phrases in domain of IT, new technologies of personal communication driving the big data new trend and internet population grew day by day but it never reach by 100%. The need of big data generated from the large companies like facebook, yahoo, Google, YouTube etc for the purpose of analysis of enormous amount of data which is in unstructured form or even in structured form. Google contains the large amount of information. So; there is the need of Big Data Analytics that is the processing of the complex and massive datasets This data is different from structured data (which is stored in relational database systems) in terms of five parameters –variety, volume, value, veracity and velocity (5V's). The five V's (volume, variety, velocity, value, veracity) are the challenges of big data management are [1]:

1. **Volume**: Data is ever-growing day by day of all types ever MB, PB, YB, ZB, KB, TB of information. The data results into large files. Excessive volume of data is main issue of storage. This main issue is resolved by reducing storage cost. Data volumes are expected to grow 50 times by 2020.
2. **Variety**: Data sources (even in the same field or in distinct) are extremely heterogeneous [1]. The files comes in various formats and of any type, it may be structured or unstructured such as text, audio, videos, log files and more. The varieties are endless, and the data enters the network without having been quantified or qualified in any way.
3. **Velocity**: The data comes at high speed. Sometimes 1 minute is too late so big data is time sensitive. Some organisations data velocity is main challenge. The social media messages and credit card transactions done in millisecond and data generated by this putting in to databases.
4. **Value**: Which addresses the need for valuation of enterprise data? It is a most important v in big data. Value is main buzz for big data because it is important for businesses, IT infrastructure system to store large amount of values in database.
5. **Veracity**: The increase in the range of values typical of a large data set. When we dealing with high volume, velocity and variety of data, the all of data are not going 100% correct, there will be dirty data. Big data and analytics technologies work with these types of data.

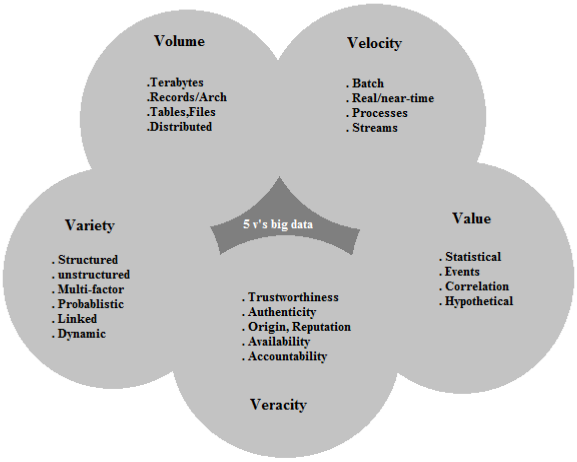


Fig. 1 Parameters of Big Data

Huge volume of data (both structured and unstructured) is management by organization, administration and governance. Unstructured data is a data that is not present in a database. Unstructured data may be text, verbal data or in another form. Textual unstructured data is like power point presentation, email messages, word documents, and instant massages. Data in another format can be.jpg images, .png images, audio files (.mp3, .wav, .aiff) and video files that can be in flash format, .mkv format or .3gp format. According to the "IDC Enterprise Disk Storage Consumption Model" report released in year 2009, in which the transactional data is proposed to raise at a composite yearly growth rate (CAGR) of 21.8%, it's far outpaced by a 61.7% CAGR calculation for unstructured data [3]. From last twenty years, the data is mounting day by day across the world in every domain. Some distinct facts about the data are, there are about 277,000 tweets per minute, 2 million queries approximately on Google every minute in all domains, 75 hours of new videos in different formats are uploaded to YouTube, More than 100 million emails are sent via Gmail, yahoo, rediff mail and many more, 350 GB of data is dealing out on facebook every day and more than 576 websites are created every minute. During the year 2012, 2.5 quintillion bytes of data were created every day. Big data and its depth analysis is the core of modern science, research area and business areas. Huge amount of data is generated from the distinct various sources either in structure or unstructured form. Such form of data stored in databases and then it become very complex to extract, transform and make in use [8]. IBM indicates that 2.5 Exabyte data is created everyday which is very difficult to analyze in various aspects. The estimation about the generated data is that till year 2003 it was represented about 5 Exabyte, then until year 2012 is 2.7 Zettabyte and till 2015 it is expected to boost up to 3 times [10].

This paper is organised as follows. In section II literature survey have been described along with advantages and disadvantages of the paper. In section III the various big data techniques has been discussed. Future Scope has been discussed in section IV for direction to emerging researchers and Final section gives a conclusion of the paper.

II. Literature Survey

1. John A. Keane [2] in 2013 proposed a framework in which big data applications can be developed. The framework consist of three stages (multiple data sources, data analysis and modelling, data organization and interpretation) and seven layers (visualisation/presentation layer, service/query/access layer, modelling/ statistical layer, processing layer, system layer, data layer/multi model) to divide big data application into blocks. The main motive of this paper is to manage and architect a massive amount of big data applications. The advantage of this paper is big data handles heterogeneous data and data sources in timely to get high performance and Framework Bridge the gap with business needs and technical realities. The disadvantage of this paper is too difficult to integrate existing data and systems.
2. Xin Luna Dong [5] in 2013 explained challenges of big data integration (schema mapping, record linkage and data fusion). These challenges are explained by using examples and techniques for data integration in addressing the new challenges raised by big data, includes volume and number of sources, velocity, variety and veracity. The advantage of this paper is identifying the data source problems to integrate existing data and systems. The disadvantage of this paper is big data integration such as integrating data from markets, integrating crowd sourcing data, providing an exploration tool for data sources.
3. Jun Wang [17] in 2013 proposed the Data-g Rouping-Aware (DRAW) data placement scheme to improve the problems like performance, efficiency, execution and latency. It could cluster many grouped data into a small number of nodes as compared to map reduce/hadoop. the three main phases of DRAW defined in this paper are: cluster the data-grouping matrix, learning data grouping information from system logs and recognizing the grouping data. The advantage of the paper is improve the throughput up to 59.8%, reduce the execution time up to 41.7% and improve the overall performance by 36.4% over the Hadoop/map reduce.
4. Yaxiong Zhao [7] in 2014 proposed data aware caching (Dache) framework that made minimum change to the original map reduce programming model to increment processing for big data applications using the map reduce model. It is a protocol, data aware cache description scheme and architecture. The advantage of this paper is, it improves the completion time of map reduce jobs.
5. Jian Tan [6] in 2013 author talks about the theoretical assumptions, that improves the performance of Hadoop/map reduce and purposed the optimal reduce task assignment schemes that minimize the fetching cost per job and performs the both simulation and real system deployment with experimental evolution. The advantage of this paper is improves the performance of large scale Hadoop clusters. The disadvantage of this paper is environmental factors such as network topologies effect on a reduce task in map reduce clusters.
6. Thuy D. Nguyen [4] (2013) author solve the multilevel secure (MLS) environmental problems of Hadoop by using security enhanced Linux (SE Linux) protocol. In which multiple sources of Hadoop applications run at different levels. This protocol is an extension of Hadoop distributed file system (HDFS). The advantage of this paper is solving environmental problems without requiring complex Hadoop server components.
7. Keith C.C. Chan [15] 2013 author describes large amount of structured and unstructured data

collection, processing and analysis from hospitals, laboratories, pharmaceutical, companies or even social media and also discus about how to collect or analyse huge volume of data for drug discovery. The advantage of this paper is how big data analytics contributes to better drug safety efficacy for pharmaceutical regulators and companies. The disadvantage of this paper it needs the algorithms that are simple, scalable, efficient and effective for data discovery process.

8. Sagiroglu, S. [8] (2013) offered the big data content, its scope, functionality, data samples, advantages and disadvantages along with challenges of big data. The critical issue in relation to the Big data is the privacy and protection. Big data samples describe the review about the environment, science and research in biological area. By this paper, we can conclude that any association in any domain having big data can take the benefit from its careful investigation for the problem solving principle. Using Knowledge Discovery from the Big data convenient to get the information from the complicated data records.

   The overall appraisal describe that the data is mounting day by day and becoming complex. The challenge is not only to gather and handle the data but also how to extract the useful information from that collected data records. In accordance to the Intel IT Center, there are several challenges related to Big Data which are rapid data growth, data infrastructure, and variety of data, visualization and data velocity.
9. Garlasu, D. [10] (2013) discussed the enhancement about the storage capabilities, the processing power along with handling technique. The Hadoop technology is widely used for the simulation purpose. Grid Computing provides the notion of distributed computing using HDFS. The benefit of Grid computing is the maximum storage capability and the high processing power. Grid Computing makes the big assistance among the scientific research and help the researcher to analyze and store the large and complex data in various formats.
10. Mukherjee, A. [11] (2012) The Big data analysis define the large amount of data to retrieve the useful information and uncover the hidden information. Big data analytics refers to the Map Reduce Framework which is discovered by the Google. Apache Hadoop is the open source platform which is used for the purpose of simulation of Map Reduce Model. In this the performance of SF-CFS is compared with the HDFS with the help of the SWIM by the facebook job traces. SWIM contains the workloads of thousands of jobs with complex and massive data arrival and computation patterns.
11. Aditya B. [12] (2012) defines big data Problem using Hadoop and Map Reduce” reports the experimental research on the Big data problems in various domains. It describe the optimal and efficient solutions using Hadoop cluster, Hadoop Distributed File System (HDFS) for storage data and Map Reduce framework for parallel processing to process massive data sets and records.

III. Big Data Technologies

Big data is a new concept for handling massive data therefore the architectural description of this technology is very new. There are the different technologies which use almost same approach i.e. to distribute the data among various local agents and reduce the load of the main server so that traffic can be avoided. There are endless articles, books and periodicals that describe Big Data from a technology perspective so we will instead focus our efforts here on setting out some basic principles and the minimum technology foundation to help relate Big Data to the broader IM domain.

A. Hadoop

Hadoop is a framework that can run applications on systems with thousands of nodes and terabytes. It distributes the file among the nodes and allows to system continue work in case of a node failure. This approach reduces the risk of catastrophic system failure. In which application is broken into smaller parts (fragments or blocks).Apache Hadoop consists of the Hadoop kernel, Hadoop distributed file system (HDFS), map reduce and related projects are zookeeper, Hbase, Apache Hive. Hadoop Distributed File System (HDFS) consists of three Components: the Name Node, Secondary Name Node and Data Node [15]. The multilevel secure (MLS) environmental problems of Hadoop by using security enhanced Linux (SE Linux) protocol. In which multiple sources of Hadoop applications run at different levels. This protocol is an extension of Hadoop distributed file system (HDFS) [12]. Hadoop is commonly used for distributed batch index building; it is desirable to optimize the index capability in near real time. Hadoop provides components for storage and analysis for large scale processing [1]. Now a day’s Hadoop used by hundreds of companies.

The advantage of Hadoop is Distributed storage & Computational capabilities, extremely scalable, optimized for high throughput, large block sizes, tolerant of software and hardware failure.

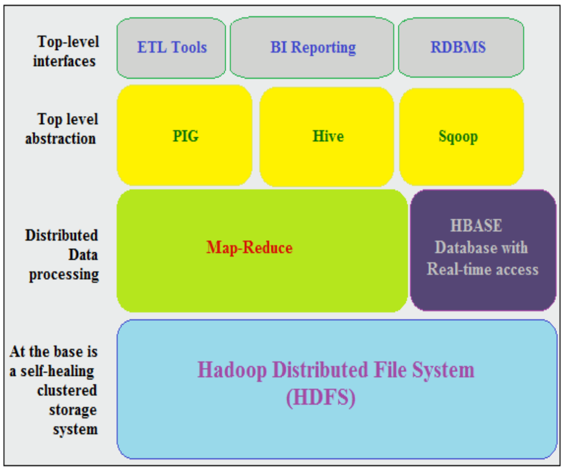

Top-level interfaces
ETL Tools
BI Reporting
RDBMS
Top level abstraction
PIG
Hive
Sqoop
Distributed Data processing
Map-Reduce
HBASE Database with Real-time access
At the base is a self-healing clustered storage system
Hadoop Distributed File System (HDFS)

Fig. 2 Architecture of Hadoop

The disadvantage of Hadoop is that it is master processes are single points of failure. Hadoop does not offer storage or network level encryption, inefficient for handling small files.

Components of Hadoop [8]:

- **HBase**: It is open source, distributed and Non-relational database system implemented in Java. It runs above the layer of HDFS. It can serve the input and output for the Map Reduce in well mannered structure.
- **Oozie**: Oozie is a web-application that runs in a java servlet. Oozie use the database to gather the information of Workflow which is a collection of actions. It manages the Hadoop jobs in a mannered way.
- **Sqoop**: Sqoop is a command-line interface application that provides platform which is used for converting data from relational databases and Hadoop or vice versa.
- **Avro**: It is a system that provides functionality of data serialization and service of data exchange. It is basically used in Apache Hadoop. These services can be used together as well as independently according the data records.
- **Chukwa:** Chukwa is a framework that is used for data collection and analysis to process and analyze the massive amount of logs. It is built on the upper layer of the HDFS and Map Reduce framework.
- **Pig**: Pig is high-level platform where the Map Reduce framework is created which is used with Hadoop platform. It is a high level data processing system where the data records are analyzed that occurs in high level language.
- **Zookeeper**: It is a centralization based service that provides distributed synchronization and provides group services along with maintenance of the configuration information and records.
- **Hive**: It is application developed for data warehouse that provides the SQL interface as well as relational model. Hive infrastructure is built on the top layer of Hadoop that help in providing conclusion, and analysis for respective queries.

B. Map Reduce

Map-Reduce was introduced by Google in order to process and store large datasets on commodity hardware. Map Reduce is a model for processing large-scale data records in clusters. The Map Reduce programming model is based on two functions which are map() function and reduce() function. Users can simulate their own processing logics having well defined map() and reduce() functions. Map function performs the task as the master node takes the input, divide into smaller sub modules and distribute into slave nodes. A slave node further divides the sub modules again that lead to the hierarchical tree structure. The slave node processes the base problem and passes the result back to the master Node. The Map Reduce system arrange together all intermediate pairs based on the intermediate keys and refer them to reduce() function for producing the final output. Reduce function works as the master node collects the results from all the sub problems and combines them together to form the output [19].

Map(in_key,in_value)--->list(out_key,intermediate_value)

Reduce(out_key,list(intermediate_value))--->list(out_value)

The parameters of map () and reduce () function is as follows:

map (k1,v1) ! list (k2,v2) and reduce (k2,list(v2)) ! list (v2)

A Map Reduce framework is based on a master-slave architecture where one master node handles a number of slave nodes [18]. Map Reduce works by first dividing the input data set into even-sized data blocks for equal load distribution. Each data block is then assigned to one slave node and is processed by a map task and result is generated. The slave node interrupts the master node when it is idle. The scheduler then assigns new tasks to the slave node. The scheduler takes data locality and resources into consideration when it disseminates data blocks.

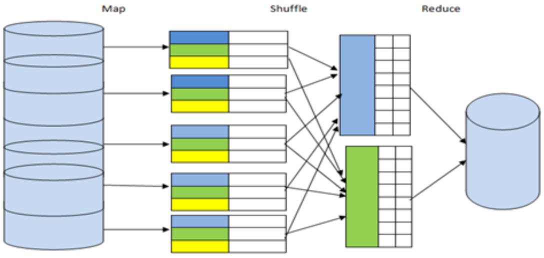


Fig. 3 Architecture of Map Reduce

Figure 3 shows the Map Reduce Architecture and Working. It always manages to allocate a local data block to a slave node. If the effort fails, the scheduler will assign a rack-local or random data block to the slave node instead of local data block. When map() function complete its task, the runtime system gather all intermediate pairs and launches a set of condense tasks to produce the final output. Large scale data processing is a difficult task, managing hundreds or thousands of processors and managing parallelization and distributed environments makes is more difficult. Map Reduce provides solution to the mentioned issues, as is supports distributed and parallel I/O scheduling, it is fault tolerant and supports scalability and it has inbuilt processes for status and monitoring of heterogeneous and large datasets as in Big Data [18]. It is way of approaching and solving a given problem. Using Map Reduce framework the efficiency and the time to retrieve the data is quite manageable. To address the volume aspect, new techniques have been proposed to enable parallel processing using Map Reduce framework [13]. Data aware caching (Dache) framework that made slight change to the original map reduce programming model and framework to enhance processing for big data applications using the map reduce model [16].

The advantage of map reduce is a large variety of problems are easily expressible as Map reduce computations and cluster of machines handle thousands of nodes and fault-tolerance.

The disadvantage of map reduce is Real-time processing, not always very easy to implement, shuffling of data, batch processing.

Map Reduce Components:

1. **Name Node**: manages HDFS metadata, doesn't deal with files directly.
2. **Data Node**: stores blocks of HDFS—default replication level for each block: 3.
3. **Job Tracker**: schedules, allocates and monitors job execution on slaves—Task Trackers.
4. **Task Tracker**: runs Map Reduce operations.

C. *Hive*

Hive is a distributed agent platform, a decentralized system for building applications by networking local system resources [8]. Apache Hive data warehousing component, an element of cloud-based Hadoop ecosystem which offers a query language called HiveQL that translates SQL-like queries into Map Reduce jobs automatically. Applications of apache hive are SQL, oracle, IBM DB2. Architecture is divided into Map-Reduce-oriented execution, Meta data information for data storage, and an execution part that receives a query from user or applications for execution.

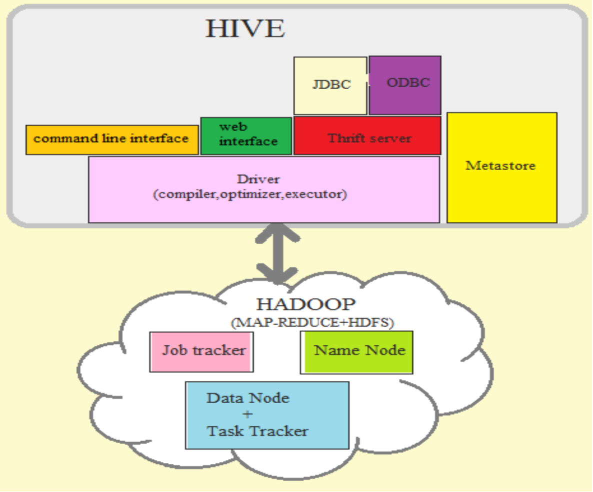


Fig. 4 Architecture of HIVE

The advantage of hive is more secure and implementations are good and well tuned.

The disadvantage of hive is only for ad hoc queries and performance is less as compared to pig.

D. *No-SQL*

No-SQL database is an approach to data management and data design that's useful for very large sets of distributed data. These databases are in general part of the real-time events that are detected in process deployed to inbound channels but can also be seen as an enabling technology following analytical capabilities such as relative search applications. These are only made feasible because of the elastic nature of the No-SQL model where the dimensionality of a query is evolved from the data in scope and domain rather than being fixed by the developer in advance. It is useful when enterprise need to access huge amount of unstructured data. There are more than one hundred No SQL approaches that specialize in management of different multimodal data types (from structured to non-structured) and with the aim to solve very specific challenges [5]. Data Scientist, Researchers and Business Analysts in specific pay more attention to agile approach that leads to prior insights into the data sets that may be concealed or constrained with a more formal development process. The most popular No-SQL database is Apache Cassandra.

The advantage of No-SQL is open source, Horizontal scalability, Easy to use, store complex data types, Very fast for adding new data and for simple operations/queries. The disadvantage of No-SQL is Immaturity, No indexing support, No ACID, Complex consistency models, Absence of standardization.

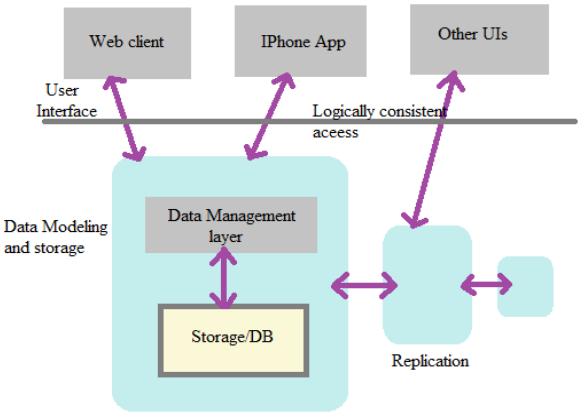


Fig. 5 Architecture of No SQL

E. HPCC

HPCC is an open source platform used for computing and that provides the service for handling of massive big data workflow. HPCC data model is defined by the user end according to the requirements. HPCC system is proposed and then further designed to manage the most complex and data-intensive analytical related problems. HPCC system is a single platform having a single architecture and a single programming language used for the data simulation.HPCC system was designed to analyze the gigantic amount of data for the purpose of solving complex problem of big data. HPCC system is based on enterprise control language which has the declarative and on-procedural nature programming language the main components of HPCC are:

- *HPCC Data Refinery*: Use parallel ETL engine mostly.
- *HPCC Data Delivery*: It is massively based on structured query engine used.
- Enterprise Control Language distributes the workload between the nodes in appropriate even load.

IV. Future Scope

There is nothing concealed that big data significantly influencing IT companies and through development new technologies only we can handle it in a managerial way. Big data totally change the way of organizations, government and academic institution by using number of tools to make the management of big data. In future Hadoop and NoSQL database will be highly in demand moving forward. The amount of data produced by organizations in next five years will be larger than last 5,000 years. In the upcoming years cloud will play the important role for private sectors and organisations to handle the big data efficiently.

V. Conclusion

In this paper we have surveyed various technologies to handle the big data and there architectures. In this paper we have also discussed the challenges of Big data (volume, variety, velocity, value, veracity) and various advantages and a disadvantage of these technologies. This paper discussed an architecture using Hadoop HDFS distributed data storage, real-time NoSQL databases, and MapReduce distributed data processing over a cluster of commodity servers. The main goal of our paper was to make a survey of various big data handling techniques those handle a massive amount of data from different sources and improves overall performance of systems.

References

[1] Yuri Demchenko “The Big Data Architecture Framework (BDAF)” Outcome of the Brainstorming Session at the University of Amsterdam 17 July 2013.

[2] Tekiner F. and Keane J.A., Systems, Man and Cybernetics (SMC), “Big Data Framework” 2013 IEEE International Conference on 13–16 Oct. 2013, 1494–1499.

[3] Margaret Rouse, April 2010 “unstructured data”.

[4] Nguyen T.D., Gondree M.A., Khosalim, J.; Irvine, “Towards a Cross Domain MapReduce Framework“ IEEE C.E. Military Communications Conference, MILCOM 2013, 1436 – 1441

[5] Dong, X.L.; Srivastava, D. Data Engineering (ICDE),” Big data integration“ IEEE International Conference on , 29(2013) 1245–1248

[6] Jian Tan; Shicong Meng; Xiaoqiao Meng; Li ZhangINFOCOM, “Improving ReduceTask data locality for sequential MapReduce” 2013 Proceedings IEEE ,1627 - 1635

[7] Yaxiong Zhao; Jie Wu INFOCOM, “Dache: A Data Aware Caching for Big-Data Applications Using the MapReduce Framework” 2013 Proceedings IEEE 2013, 35 - 39 (Volume 19)

[8] Sagiroglu, S.; Sinanc, D.,”Big Data: A Review”,2013,20-24

[9] Minar, N.; Gray, M.; Roup, O.; Krikorian, R.; Maes, “Hive: distributed agents for networking things“ IEEE CONFERENCE PUBLICATIONS 1999 (118-129)

[10] Garlasu, D.; Sandulescu, V.; Halcu, I.; Neculoiu, G,”A Big Data implementation based on Grid Computing”, Grid Computing, 2013, 17-19

[11] Mukherjee, A.; Datta, J.; Jorapur, R.; Singhvi, R.; Haloi, S.; Akram, “Shared disk big data analytics with Apache Hadoop”, 2012, 18-22

[12] Aditya B. Patel, Manashvi Birla, Ushma Nair, “Addressing Big Data Problem Using Hadoop and Map Reduce”, 2012, 6-8

[13] Jefry Dean and Sanjay Ghemwat, MapReduce:A Flexible Data Processing Tool, Communications of the ACM, Volume 53, Issuse.1,2010, 72-77.

[14] Chan,K.C.C. Bioinformatics and Biomedicine (BIBM), “Big data analytics for drug discovery” IEEE International Conference on Bioinformatics and Biomedicine 2013,1.

[15] Kyuseok Shim, MapReduce Algorithms for Big Data Analysis, DNIS 2013, LNCS 7813, pp. 44–48, 2013.

[16] Wang, J.; Xiao, Q.; Yin, J.; Shang, P. Magnetics, “DRAW: A New Data-gRouping-AWare Data Placement Scheme for Data Intensive Applications With Interest Locality“IEEE Transactions (Vol: 49), 2013, 2514 – 2520

[17] HADOOP-3759: Provide ability to run memory intensive jobs without affecting other running tasks on the nodes.

Comparative Study of Conventional Random Diagonal Code and Random Diagonal Code with EDFA for Spectral Amplitude-Coding Optical CDMA System

Bhanu Priya[1] and Himali Sarangal[2]
[1,2]*Department of Electronics and Communication Engineering,*
GNDU Regional Campus Jalandhar, Punjab, India
E-mail: [1]bpriya812@gmail.com, [2]himali.sarangal@gmail.com

***Abstract*—In this paper we present and compare the performance of random diagonal (RD) code for spectral-amplitude coding OCDMA (SAC-OCDMA) system using a newly proposed spectral direct detection technique and with the use of EDFA. By comparing the theoretical and simulation results taken from the commercial optical systems simulator "Optisystem 7.0 " we show that the proposed new spectral direct detection technique utilizing RD code with the collaboration of EDFA considerably improves the performance compared with the conventional Random diagonal codes.**

Keywords: Multiple Access Interference, Random Diagonal Codes

I. Introduction

The concept of the OCDMA systems is the data extraction by a user in the presence of other users or in the presence of multiple access interference (MAI). MAI is the dominant source of corruption in an OCDMA system, therefore efficient design of the code sequences and detection scheme is required to reduce the affect of MAI [1,2]. The presence of multiple users utilizing the same medium, at the same instant and frequencies to transmit their data streams parallel in the OCDMA systems produce MAI. It has been assumed that the cross-correlation value is always equal to zero, because of the data signal is present only on the data segment which further makes the noise equal to zero. In this paper, a new version of code that is RD code has been proposed. It has been assumed that the cross-correlation value is variable, and the data is out-of-phase autocorrelation and cross-correlation values are used [3]. This paper is organized as follows. In section II we will discuss how the code is been develop theoretically and its properties. In section III, we focus on performance analysis of the new code and finally the result and conclusion in section IV & V.

II. Review of Random Diagonal Code

(N, W, λ) is the representation of the random diagonal code where N is the code length, W is the code weight, and λ is it in-phase cross correlation. The blueprint of RD code can be performed by dividing the code sequence into two groups which is code segment and data segment. [4]

A. Data Segment

This segment is represented by a matrix containing only one "1" to keep cross correlation zero at data level ($\lambda = 0$). To understand this concept we can use the matrix ($K \times K$) where K will represent number of users. These matrices contain binary coefficient and a basic Zero cross code (weight = 1) which is defined by $[Y1]$. $Y1$ can be expressed as

$$[\mathrm{Y1}] = \begin{bmatrix} 100 \\ 010 \\ 001 \end{bmatrix}$$

where $[Y1]$ consists of ($K \times K$) identity matrices. Notice, for the above expression the cross correlation between any two rows is always zero.

A good set of code is used to obtain the maximum number of codes with maximum weight and with Erbium Doped Fiber Amplifier is significantly better minimum length with the best possible autocorrelation and simulations that the transmission performance of the RD code cross-correlation properties. A code with a larger size should respectively. It will be shown by theoretical studies and give better BER performance. Hence unipolar {0,1} codes which present out on the data segment, and code segment have peculiar property described

B. Code Segment

The code segment matrix can be expressed as follows for $W = 4$, Where $[Y2]$ consists of two parts—weight matrix part $[W]$ and basics matrix part $[B]$. Basic part $[B]$ can be expressed as

$$[\mathrm{Y2}] = \begin{bmatrix} 11010 \\ 01101 \\ 10110 \end{bmatrix}$$

$$[\mathrm{B}] = \begin{bmatrix} 011 \\ 110 \\ 101 \end{bmatrix}$$

And weight part which called M matrix

$$[\mathrm{M}] = \begin{bmatrix} 10 \\ 01 \\ 10 \end{bmatrix}$$

which is meant for incrementing number of weights, let Mi = [M1 M2 M3 …. Mi], where i represents number of [M] matrix and i, is given by

$$I = W - 3 \qquad (1)$$

For Lth user matrix [*M*] and [*B*] can where *j* represents the value for Nth user (*j* = 1*,* 2*, . . .* ,*L*), and the value of *a j* is either zero or one [5].The weights for code part for both matrix [*M*], [*B*] are equal to *W*−1, so the total combination of code is represented as (L × *N*) where *L* = 3, *N* = 8, as given by [*Z*1],

$$[Z1] = [Y1|Y2]\ [Z1] = \begin{bmatrix} 00111010 \\ 10100110 \\ 10010110 \end{bmatrix}$$

for example if W=5, from Eq.(1) i = 2, so that M2= [M1 M2]

$$[M2] = \begin{bmatrix} 1010 \\ 0101 \\ 1010 \end{bmatrix}$$

Notice that to increase the number of users simultaneously with the increase of code be expressed as

$$M(j) = \begin{bmatrix} 01 \\ 10 \\ 01 \\ 10 \\ .. \\ .. \\ .. \\ .. \\ .. \end{bmatrix} \qquad B(j) = \begin{bmatrix} 011 \\ 110 \\ 101 \\ 011 \\ .. \\ .. \\ .. \\ .. \\ .. \end{bmatrix}$$

III. Performance Analysis

The setup of the conventional RD system using spectral direct detection technique with three users is shown in Fig. 1. Each chip has a spectral width of 0.8 nm. The tests were carried out at a rate of 10 Gb/s for 20-km distance with the ITU-T G.652 standard single-mode optical fiber (SMF).All the attenuation *α* (i.e., 0.25 dB/km), dispersion (i.e., 18 ps/nm km), and nonlinear effects were added. After the splitter, we used a fiber Bragg grating (FBG) spectral amplitude decoder operates to decode the data at data sub-matrix [5].

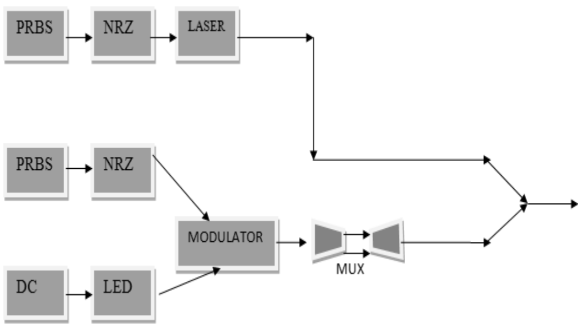


Fig. 1 RD Segment

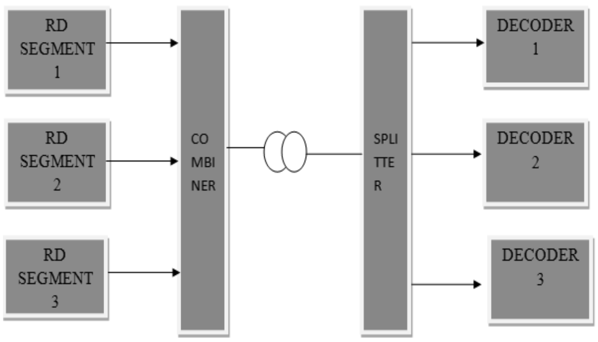


Fig. 2 Simulation Setup of Conventional RD Code without EDFA

The signal was decoded by a photo-detector (PD) followed by low-pass filter (LPF) and error detector, respectively. The transmitted power used was − 10 dBm out of the broadband source. The noise generated at the receivers was set to be random and uncorrelated. The dark current value was 5 nA, and the thermal noise coefficient was 1.8 × 10−23 W/Hz for each of the photo detectors. Fig. 3 shows RD code with EDFA with same parameters and we come to know that RD code performance get enhanced with use of Erbium Doped Fiber.

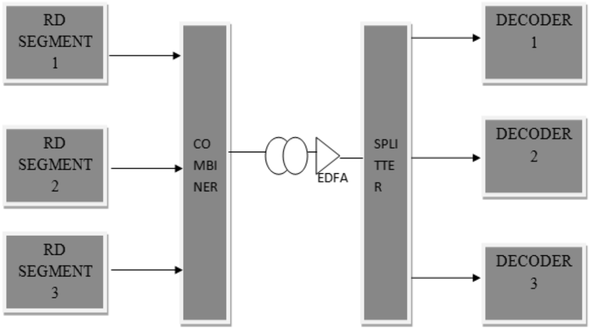
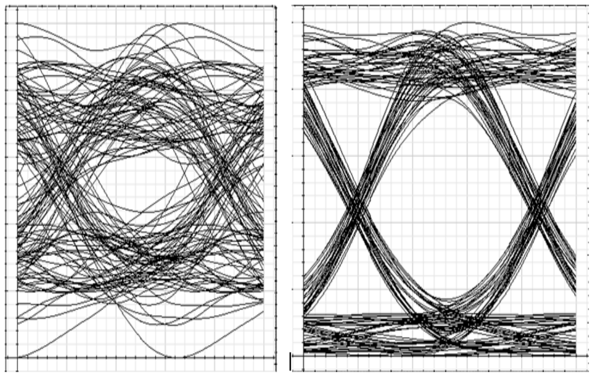


Fig. 3 Simulation Setup of Conventional RD Code with EDFA

IV. Results

The eye pattern diagrams for RD and RD with EDFA codes are shown in Fig. 3. The eyes diagram shown in Fig. 3 clearly depict that the RD code using EDFA gives better performance, having a larger eye opening. The corresponding simulated BER for RD and RD with EDFA codes systems are shown in Fig. 4. The vertical distance between the top of the eye opening and maximum signal level gives the degree of distortion. The Bit Error Rate and eye height of conventional RD code is 0.0017395 and 3.55e-008 and that with EDFA is 2.88348*10exp (-16) and0.00245339 respectively. So this comparison clears the fact that the performance of RD code with EDFA is best. By studying both the Eye patterns we can clearly observe that the Fig is completely scattered and the latter is open wide.



(a) (b)

Fig. 4 Eye Diagram of (a) One of the RD Channels (b) One of the RD Channels with EDFA, at 10 Gbit/s

V. Conclusion

In this paper, we have studied the comparison of Conventional Random Diagonal code and Random Diagonal Code with EDFA for Spectral Amplitude-Coding Optical CDMA System. It has been shown that through simulation results RD code with EDFA get better BER performance than conventional one and improved the overall system performance, so as a result RD code with EDFA can also be used in synchronous optical CDMA system for the cancellation of MUI. After simulation we realized that RD codes are shorter and have zero cross correlation which makes phase induced intensity noise zero. Further research on RD Code with EDFA at 20 Gigabits/second can be carried out. [4].

References

[1] Lei Xu, I. Glesk, V. Baby, P.R. Prucnal, Multiple access interference (MAI) noise reduction in a 2D optical CDMA system using ultrafast optical thresholding,in: 17th Annual Meeting of the IEEE, vol. 2, November 8–9, 2004, Lasers and Electro-Optics Society, 2004, pp. 591–592.

[2] Jen-Fa Huang, Chao-Chin Yang, Reductions of multiple-access interference in fiber-grating-based optical CDMA network, IEEE Trans. Commun. 50 (10) (2002) 1680–1687.

[3] R.C. Dixon, Spread Spectrum Systems with Commercial Applications, John Wiley & Sons, New York, 1994.[4] HilalFadhil, Syed Aljunid, and Badlished Ahmed,Performance of OCDMA Systems Using Random Diagonal Code for Different Decoders Architecture Schemes,The International Arab Journal of Information Technology, Vol. 7, No. 1, January 2010.

[4] HilalAdnanFadhil ,S.A.Aljunid, R.B. Ahmad,Performance of random diagonal code for OCDMA systems using new spectral direct detection technique,Optical Fiber Technology 15 (2009) 283–289.

[5] J.A. Salehi, Code division multiple-access techniques in optical fiber networks.Part I. Fundamental principles, IEEE Trans. Commun. 37 (8) (1989)824–833.

[6] J. Shah, Optical CDMA, Opt. Photon. News 14 (2003) 42–47.

[7] I.B. Djordjevic, B. Vasic, Combinatorial constructions of optical orthogonalcodes for OCDMA systems, IEEE Commun. Lett.8 (6) (2004).

[8] H. Yin, Optical Code Division Multiple Access Communication Networks Theoryand Application, Springer Link, 2009.

[9] H.A.Fadhil, S.A.Aljunid, R.B.Ahmad, Design considerations of high performance optical CDMA: a new spectral amplitude code based on laser and LED lightsource IET Optoelectronics Journal, vol. 4, The Institution of Engineering and Technology, UK, 2010, pp. 29–34.

Performance Evaluation of CI Based Routing Protocols for WSNs

Palvinder Singh Mann[1] and Satvir Singh[2]
[1]*Department of Computer Sc. & Engineering,*
DAV Institute of Engineering & Technology, Jalandhar–144008, India
[2]*Department of Electronics & Communication Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab–152004, India
E-mail: [1]psmaan@hotmail.com, [2]drsatvir.in@gmail.com

***Abstract*—Recent advances in Wireless Sensor Networks (WSNs) have led to new paradigm for protocol design especially for sensor networks where energy awareness is an essential component. Most of the research, however, has been focused on to the development of routing protocols since they might differ depending on the application area and network architecture. Computational Intelligence (CI) based optimization techniques paved the way for energy efficient routing for WSNs. In this paper, we evaluated and analyse the performance of CI based routing protocols with classical LEACH protocol. Our simulation results shows that CI based routing protocol perform better in terms of energy consumption thus increasing network life over classical protocol.**

Keywords: Wireless Sensor Networks, Computational Intelligence, Energy-efficient Routing

I. INTRODUCTION

Wireless Sensor Networks (WSNs) consist of large number of sensor nodes which sense, process and transmit data from an area of deployment. Sensor nodes should send the collected data to a Base Station (BS) or Sink for further processing and analysis. A Wireless Sensor Network design must have low power consumption with flexible scalability and robust network adaptability coupled with good fault-tolerance [3]. The key area to achieve the above mentioned features lies in efficient data routing between sensor nodes and base station. Many researchers are currently engaged in developing routing protocols that fulfill the requirements of these key features. The main aim is to find ways for energy-efficient route setup and reliable relaying of data from the sensor nodes to the base station so that the lifetime of the network is maximized [5]. There is always a trade off between computation and communication in each node when it makes the route decision and data aggregation. As the size of WSNs grows, so does the complexity of the data routing. Therefore a key area of WSNs research is in developing new routing algorithms to meet the strict low-power limitations. Computational Intelligence (CI) based routing protocols presents an optimized solutions to the energy constrained WSNs over the classical approaches. Below we discuss some classical and CI based approaches for WSNs Routing.

II. DESIGN CHALLENGES FOR WSN'S ROUTING

Designing of routing protocols for WSNs is a challenging task due to the following parameters:

A. *Minimal Computational Requirements*

Sensor nodes are typically equipped with a low-end CPU and have limited memory. Therefore, it is customary that the routing algorithm has minimal processing overhead to make its execution feasible and effective on such a low-end processor.

B. *Energy Efficiency*

Sensor nodes are equipped with small non-rechargeable batteries therefore; the efficient battery utilization of a sensor node is a critical aspect to support the extended operational lifetime of the individual nodes and of the whole network. A WSN routing protocol is expected to:

1. Minimize the total number of transmissions involved in route discovery and data delivery.
2. Distribute the forwarding of the data packets across multiple paths, so that all nodes can deplete their batteries at a comparable rate. This will result in the overall increase of the network lifetime.

C. *Self-organization*

A WSN is expected to remain operational for an extended period of time. During this time, new nodes might be added to the network, while other nodes might incur in failures or exhaust their batteries, becoming un operational. A routing protocol must be resilient to such dynamic and generally unpredictable variations and must sustain the long-term availability of essential network services. Therefore, the network protocols, and the routing protocols in particular, must be empowered with self-organizing and self-management properties.

D. *Scalability*

In a wide range of WSN applications, thousands or even millions of nodes are expected to be deployed. A typical example is battle field surveillance, in which the criticality and the geographical extension of the scenarios require the deployment of large numbers of densely distributed sensors that have short communication ranges and high failure rates. Therefore, the routing protocol should be able to effectively cope with the challenges deriving from intensive radio interference, very long paths, and unpredictable failures. Moreover, it should be able to display scalable performance in face of these challenges.

E. *Data Aggregation*

Sensor networks can generate large amounts of locally redundant data. For instance, when a node detects that the temperature in its surroundings has exceeded a certain threshold value, it is likely that also its neighbouring nodes will detect the same event. If all these sensor nodes notify the event to the monitor node, which then can aggregate the received information to assess the event with high statistical confidence. The downside of this way of proceeding lies in the excessive use of network resources. However, not every single piece of information need to be communicated to the global sink. Information from a group of neighbouring nodes can be partially aggregated and processed as close as possible to its origin. In this way, it is possible to significantly reduce the number of transmissions, saving on the limited available hardware resources and reducing the negative effects due to radio interference. A good routing protocol for WSNs must be able to effectively support the setup and the use of data paths for in-network data aggregation.

III. Taxonomy of WSN's Routing

Depending upon the requirements of the application and area of deployment WSNs routing protocols are divided on various taxonomies.

A. *Data-centric Protocols*

Data-centric protocols will combine the application needed to access data with a natural framework for in-network processing. In many applications of WSNs, due to lack of global identification along with random deployment of sensor nodes, it is hard to select a specific set of sensor nodes to be queried. This consideration has led to data-centric routing, which is different from traditional address-based routing where routes are created between addressable nodes. In data-centric routing, the sink sends queries to certain regions and waits for data from the sensors located in the selected regions. Since data is being requested through queries, attribute-based naming is necessary to specify the properties of data.

B. *Location Based Protocols*

The idea of location-based protocols is using an area instead of a node identifier as the target of a packet. Any node which is positioned within the given area will be acceptable as a destination node and can receive and process a message. In the context of sensor networks, such location-based routing is evidently important to request sensor data from some region. Since there is no addressing scheme for sensor networks like IP-addresses and they are spatially deployed in a region, location information can be utilized in routing data in an energy-efficient way. For instance, if the region to be sensed is known, using the location of sensor nodes, the query can be number of transmission significantly.

C. *Hierarchical Protocols*

The main aim of hierarchical routing is to efficiently maintain the energy consumption of sensor nodes by involving them in multi-hop communication within a particular cluster and by performing data aggregation and fusion in order to decrease the number of transmitted messages to the sink. Cluster formation is typically based on the energy reserve of sensors and sensor's proximity to the cluster head. LEACH is one of the first hierarchical routing approaches for sensors networks.

IV. Approaches for WSN's Routing

Researchers around the world have developed Classical as well as CI based routing protocols for WSNs.

A. *Classical Based Routing Protocols*

Some of the most popular classical routing protocols are discussed below, which have addressed some of the most challenging aspects of the WSN's routing.

1. *Directed Diffusion (DD):* Taxonomy—Data Centric in Directed Diffusion (DD) events are diffused through sensor nodes by using a naming scheme for it. Attribute value pairs for the event is adopted while querying the sensors in an on demand basis. It is a popular data aggregation paradigm for WSNs. It is a data-centric and application aware paradigm in the sense that all data generated by the sensor nodes is named by attribute-value pairs. Creation of query is achieved by defining an interest using a list of attribute value pairs such as name of objects, duration of the event, and geographical location etc. DD is specific to some applications of sensor networks due to its query-driven data delivery model, since those requiring continuous data delivery to the sink will not perform efficiently.
2. *Sensor Protocol for Information via Negotiation (SPIN):* Taxonomy—Data Centric in SPIN [18], three messages are defined to aid in data dissemination: ADV message for advertisement of data, REQ message for data request, and DATA message that carry the actual information. In SPIN, data are named using meta-data. The protocol meta-data negotiation helps in elimination of overlapping, redundant information and resource blindness. The advertisement method of SPIN does not guarantee the delivery of data as nodes that are interested in the data may be far

away from the source node, and nodes in between the source and the sink may not be interested. In that case, such data will not get to the base station.

3. *Geographic and Energy-Aware Routing (GEAR):* Taxonomy—Location Based in GEAR [19], each sensor node is equipped with a GPS sensor for location identification. The protocol utilizes energy aware heuristics which is based on geographic information for the selection of nodes to route data to the sink, and uses geographically recursive forwarding algorithm for data dissemination within the target area. The main idea is to restrict the number of interests in directed diffusion by only considering a certain region rather than sending the interests to the whole network. By doing this, GEAR can conserve more energy than directed diffusion and proves to be a energy efficient routing protocol but GPS device add extraordinary cost to sensor [3].
4. *Low-Energy Adaptive Clustering Hierarchy (LEACH):* Taxonomy—Hierarchical Protocol LEACH [20] became the most popular and the first energy-efficient hierarchical algorithm proposed for power consumption reduction in sensor networks. LEACH rotates the clustering task among the participating nodes based on duration. Each cluster head communicates directly to the sink. The algorithm is also based on data aggregation or fusion techniques as the original data is combined and aggregated into smaller size of data that carry only required information to all individual nodes. Cluster heads change randomly over time so as to balance the energy dissipation of nodes. The protocol is completely distributed and requires no global knowledge of the network. As it uses formation of cluster heads, or dynamic clustering, it brings extra overheads, hence diminishing the gain in energy saving.
5. *Power-Efficient Gathering in Sensor Information Systems (PEGASIS):* Taxonomy—Hierarchical Protocol PEGASIS [21] is an improved version of LEACH. It avoids the formation of multiple clusters. Each node can transmit and receive data from a neighbour and only one node is selected from a chain at a time to communicate with the sink. Data is combined and moved from node to node, aggregated and sent to the sink. However, the protocol introduces excessive delay for distant nodes on the chain. In addition, the single leader exhausts its energy as it involves regular transmission.
6. *Energy-aware QoS Routing Protocol (EAQSR):* Taxonomy—QoS Based Energy aware QoS routing [23] is a table driven multi-path routing protocol with embedded QoS in its routing decision. Its aim is to find an optimal path to the gateway in terms of energy consumption and error rate while meeting the end-to-end delay requirements. Both the paths that meet the requirements for real-time traffic, as well as well as maximizing the throughput for non-real time traffic were considered.
7. *A Stateless Protocol for Real-Time Communication in Sensor Networks (SPEED):* Taxonomy—QoS Based SPEED [24] is a QoS routing protocol for sensor networks. The protocol involves three types of communication techniques: real-time unicast, real-time area-multicast and real-time area-anycast. It requires each node to maintain information about its neighbours and uses geographic forwarding in order to locate the paths. The protocol is aimed to be a stateless and localized algorithm with minimal control overhead. The protocol provides end-to-end soft real-time communication by maintaining a desired delivery speed across the sensor network through a novel combination of feedback control and nondeterministic geographic forwarding. SPEED is a highly efficient and scalable protocol for sensor networks where the resources of each node are scarce.

B. *Computational Intelligence (CI) Based Routing Protocols for WSN's*

Routing in WSN's remain a challenge for researchers as various classical protocols lacks on energy efficiency, fault tolerance or on scalability. Researchers around the world have developed some robust protocols based on Computational Intelligence (CI), which provide optimal solutions to the above mentioned problems. Some of the CI based routing protocols are listed below:

1. *Pheromone Based Energy Aware Directed Diffusion (PEADD):* Taxonomy—Data Centric PEADD [7] is a variant of DD, based on ant colony optimization heuristic. The protocol is aimed at maximizing the lifetime of the sensor networks by involving nodes with higher energy in the information gathering process. In this algorithm ants increase the pheromone on a path proportionally to the remaining energy levels of the nodes. Paths with larger residual energy are increased, while others are reduced i.e. the amount of pheromone decay with transmitting data because the pheromone is linked to the remaining energy. The pheromone level is updated based on the amount of transmitting data. The algorithm use the same route selection and updating as that of the general ant based routing as described above [3].
2. *Comprehensive Routing Protocol (CRP):* Taxonomy—Data Centric CRP [8] algorithm is an improved version of energy aware routing (EAR) and based on ant colony optimization, but in its

routing decision, it uses probability of selection of which it considers the network lifetime and data packet arrival rate. The protocol argues that always using the path which is considered as the best and optimal path from the point of view might not be the best as it will lead to depletion of the path nodes energy and instead proposes the use of sub-optimal paths occasionally. The protocol has three phases: routing table setup, data communication, and route maintenance.

3. *Sensor Driven and Cost-Aware Ant Routing (SC):* Taxonomy—Location Based in SC [9], it is assumed that ants have sensors so that they can smell where there is food at the beginning of the routing process so as to increase in sensing the best direction that the ant will go initially. In addition to the sensing ability, each node stores the probability distribution and the estimates of the cost of destination from each of its neighbours. It suffers from misleading data when there is obstacle which might cause errors in sensing.
4. *Ant Colony Clustering Algorithm (ACALEACH):* Taxonomy—Hierarchical Protocol ACLEACH [11] is based on Ant Colony Clustering Algorithm, which is an ant colony based improved version of LEACH. The algorithm not only considers the node residual energy, but also the distance between the cluster heads was considered in selection of cluster heads. It applies the ACA into inter-cluster routing mechanism to reduce the energy consumption of cluster heads and finally prolong the lifetime of sensor networks. The protocol did not consider throughput and delay in its routing process, and hence may also be weak in energy efficiency due to overheads.
5. *Ant Colony Based Multipath Routing Algorithm (ACMRA):* Taxonomy—Hierarchical Protocol ACMRA [12] discover disjoint multipath between the source nodes and sink node. In multipath routing, multiple paths between source and destination are established. The algorithm generates two types of ants: search ant (SANT) and reinforcement ant (RANT). SANT is used to collect information about paths and the intermediate nodes local information as they travel along the path. RANT is used to update the pheromone table along the reverse path, and bring information of path to source node, such as residual energy of node, path length and energy consumption of the current path. It is an on demand multipath protocol and adopts a two-phase routing process involving the constructing routing and data transmission phases. In the constructing routing phase, cluster head in the event region generates SANTs according to the number of neighbour nodes, and chooses the next node to move to according to probability of selection. While in the data transmission phase, the network lifetime relates to hop count, energy consumption and the minimum energy at a path [3].
6. *Energy-Aware Evolutionary Routing Protocol (ERP):* Taxonomy—Hierarchical Protocol Energy-Aware Evolutionary Routing Protocol (ERP) guarantee better tradeoff between lifespan and stability period of a network with efficient energy utilization, is based on evolutionary algorithms(EAs).
7. *Multipath Routing Protocol (MRP):* Taxonomy—Hierarchical Protocol Multipath Routing Protocol (MRP), is based on dynamic clustering and CI based ant colony optimization (ACO). A CH is selected among nodes located in the event area and an improved ACO algorithm is applied in the search for multiple paths between the CH and sink node. MRP prolonged the network lifetime and reduces the average energy consumption effectively.
8. *Energy Efficient Ant Based Routing (EEABR)*: Taxonomy—QoS Based EEABR [13] is based on Ant Colony Optimization (ACO) metaheuristic. In this protocol, each node in the network launches a forward ant at a regular interval with the aim of finding a route to the destination (sink). In the protocol, each ant only carries the address of the last visited nodes which means intermediate nodes carries the records of received and forwarded ants in the tables. The table content of each node contains the previous node, forward node, ant identification, and timeout value. Each time a node receive a forward ant, it looks up its table to search for any possible loop. If no loop exists, the node saves into its table the information of the ant and restarts a timer and forwards it to the next hop. When the forward ant reaches its destination, it is converted to backward ant with the mission to update the pheromone trail of the path traversed by the forward ant.
9. *Bee-Inspired Power Aware Routing (Beesensor):* Taxonomy—QoS Based Beesensor [16] is an algorithm based on the foraging principles of honey bees with an on-demand route discovery (AODV). The algorithm works with three type of agents; packers, scouts and foragers. Packers locate appropriate foragers for the data packets at the source node. Scouts are responsible for discovering the path to a new destination using the broadcasting principle. Foragers are the main workers of BeeSensor which follow a point-to-point mode of transmission and carry the data packets to a sink node. When a source node detects an.

V. Simulation and Analysis

We simulated well known classical routing protocol LEACH, CI based hierarchal routing protocols MRP and ERP on *Nature Inspired Tool for Sensor Simulation* (NITSS), a java based open platform developed to evaluate the performance of WSNs routing protocols. We have evaluated the performance of these protocols based on following performance parameters:

A. *Packet Delivery Ratio (PDR)*

It is the ratio of total number of packets received at BS to the total number of packets generated at all sensor nodes.PDR gives the percentage of total packets delivered at BS. This ratio is vital as it reflects the robust performance of the protocol. Simulation results of PDR are shown in Table below:

Table 1 Packet Delivery Ratio

No. of Nodes	MRP	LEACH	ERP
30	99.95092	98.40863	99.76959
75	99.97161	99.90158	99.95242
115	99.98433	99.95214	99.93916
175	99.98753	99.98624	99.97282
220	99.97762	99.92258	99.96794

Performance of MRP is highest for delivering packets as compared to ERP and LEACH (Fig. 1.) It is evident from Fig. 1. that packet delivery ratio of MRP remains nearly 100 % even when number of nodes increased to 220. The performance of LEACH is comparable to ERP which also delivers almost 99% packets. But it is important to note that CI based routing protocols outperform over classical protocols in terms of percentage of packet delivery.

We also evaluated other performance parameters like average energy consumed and network life for WSNs. CI based routing protocols paved the way for energy efficient solutions for variety of applications and wide range of WSNs which incorporate the principals of metaheuristic to solve optimization problems like routing for WSNs which is very vital for the network performance.

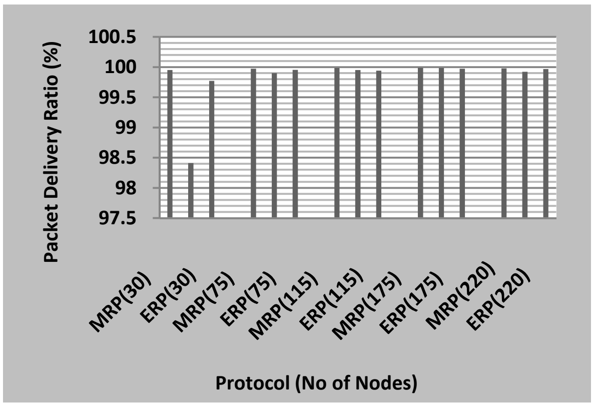


Fig. 1 Packet Delivery Ratio

B. *Average Energy Consumed*

It is the average energy consumed by all the sensor nodes in delivering all data packets generated at all its nodes to BS. This is another most significant parameter to determine the energy efficiency of the protocol. The below Table shows the results of average energy consumed at various nodes.

Table 2 Average Energy Consumed

No. of Nodes	MRP	LEACH	ERP
30	0.00102	0.00105	0.00109
75	0.00116	0.00124	0.00122
115	0.00136	0.0015	0.00131
175	0.00127	0.0014	0.00105
220	0.0011	0.00138	0.00089

MRP consumes less energy initially as compared to LEACH and MRP (Fig. 2.) when number to nodes are less but gradually when number of nodes increased, ERP perform better in terms of less amount of energy consumption.

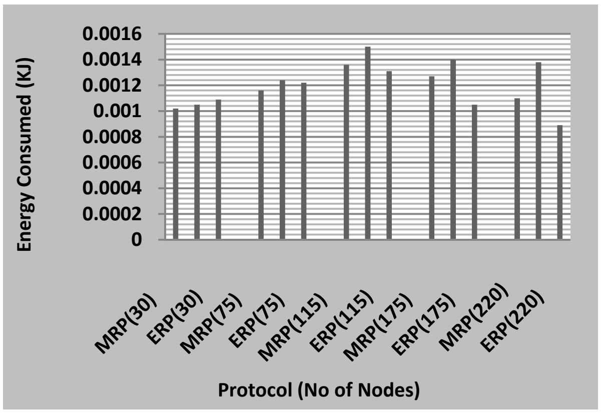


Fig. 2 Average Energy Consumed

C. *Network Life*

Network Life is one of the most significant parameter which describes the extended life of the network. If a protocol supports more number of rounds, it ultimately increases the life of the network; hence more packets can be delivered at BS for longer duration. The below Table, shows the comparative results for number of rounds of LEACH, ERP and MRP.

Table 3 Number of Rounds (Network Life)

No. of Nodes	MRP	LEACH	ERP
30	293	257	175
75	290	247	195
115	278	229	207
175	235	195	215
220	218	168	209

In terms of Network Life, which is most vital parameter of network performance, it is evident form

Fig 3. that MRP took largest number of rounds as compared to ERP and LEACH.

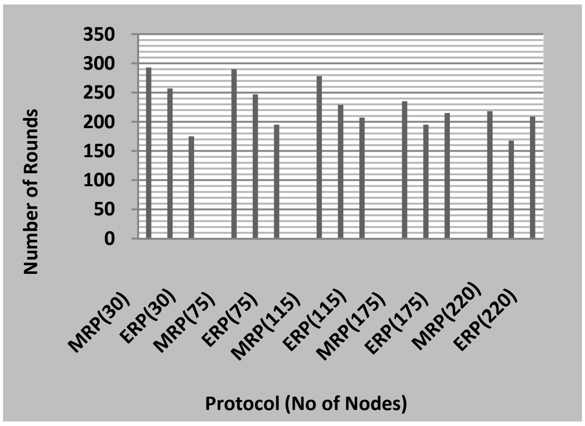


Fig. 3 Network Life

Another important factor is that CI based routing protocols provide increased network life to WSNs.

VI. Conclusion

WSNs consist of hundreds of thousands of nodes which collect vital data for wide range of applications. Due to constraints like less battery power, routing remains a challenge for researchers around the world. CI based routing protocols paved the way for efficient data routing for WSNs and our evaluation proved that CI based routing protocols outperformed over classical approaches. It encourages us to incorporate other nature inspired CI based methods to build efficient protocols for WSNs.

References

[1] Waltenegus Dargie and Christian Poellabauer "Fundamentals of Wireless Sensor Networks: Theory and Practice" Xuemin (Sherman) Shen et. al (eds.), *Wireless Communications and Mobile Computing,* John Wiley & Sons Ltd., 2010.

[2] Peng Jiang, Yu Wen, Jianzhong Wang, Xingfa Shen and Anke Xue "A Study of Routing Protocols in Wireless Sensor Networks" *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, June 21 - 23, 2006.

[3] Adamu Murtala Zungeru, Li-Minn Ang and Kah Phooi Seng "Classical and Swarm Intelligence Routing Protocols for Wireless Sensor Networks: A Survey and Comparison" *Journal of Network and Computer Applications, Elsevier,* 2012.

[4] Muhammad Saleem, Gianni A. Di Caro and Muddassar Farooq "Swarm intelligence based routing protocol for wireless sensor networks Survey and future directions" *Information Sciences, Elsevier,* Vol. 181, pp 4597–4624, 2011.

[5] Kemal Akkaya and Mohamed Younis "A survey on routing protocols for wireless sensor networks" *Ad Hoc Networks, Elsevier,* Vol. 3, pp 325–349, 2005.

[6] Jamal N. AL-Karaki and Ahmed E. Kamal "Routing Techniques in Wireless Sensor Networks: a Survey" *IEEE Wireless Communications,* pp 6-28, December 2004.

[7] X. Zhu "Pheromone based energy aware directed diffusion algorithm for wireless sensor network", *In Proceedings of the International Conference on Intelligent Computing* (ICIC),2007.

[8] W Guo,W Zhang and G Lu " A Comprehensive Routing Protocol in Wireless Sensor Network Based on Ant Colony Algorithm" *In Proceedings of the Second International Conference on Networks Security, Wireless Communications and Trusted Computing,* 2010.

[9] Y Zhang, LD Kuhn and MPJ Fromherz "Improvements on Ant Routing for Sensor Networks*" Ant Colony Optimization and Swarm intelligence, Lecture Notes Computer Science,* pp 289-313, 2004.

[10] L Wang, R Zhang and A N Model " An Energy-Balanced Ant-Based Routing Protocol for Wireless Sensor Networks", *Network,* pp 1-4, 2009.

[11] G Wang,Y Wang and X Tao " An Ant Colony Clustering Routing Algorithm for Wireless Sensor Networks" *In Third International Conference on Genetic and Evolutionary Computing,* pp 670-73, 2009.

[12] J Yang, Y Lin, W Xiong and B Xu "Ant Colony-Based Multipath Routing Algorithm for Wireless Sensor Networks" *International Conference on Computer Science and Software Engineering,* pp 2-5, 2008.

[13] T Camilo, C Carreto, J S Silva and F Boavida "An Energy Efficient Ant-Based Routing Algorithm for Wireless Sensor Networks" *Ant colony Optimization and Swarm Intelligence*, pp 49- 59,2006.

[14] Wen Y-feng, Chen Y-quan and Pan M " Adaptive ant-based routing in wireless sensor networks using Energy Delay metrics*", Journal of Zhejiang University* ; Vol. 9, pp 531-538,2008.

[15] R Ghasemaghaei, MA Rahman, W Gueaieb and A Saddik " Ant Colony-Based Reinforcement Learning Algorithm for Routing in Wireless Sensor Networks" *IEEE Instrumentation & Measurement Technology Conference IMTC,* pp 1-6, 2007.

[16] M Saleem and M Farooq " Beesensor: A bee-inspired power aware routing protocol for wireless sensor Network" *In Proceedings of EvoWorkshops (EvoCOMNET),*pp 81-90,2007.

[17] C Intanagonwiwat, R Govindan,D Estrin,J Heidemann and F. Silva " Directed diffusion for wireless sensor networking" *In proceedings of IEEE/ACM Transactions on Networking,* Vol. 11,pp 2-16, 2003.

[18] J Kulik, W Heinzelman, H Balakrishnan " Negotiationbased protocols for disseminating information in wireless sensor networks",*Wireless Networks,* Vol 8, pp 169-85, 2002.

[19] Y Yu, R Govindan and D Estrin " Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks*" In Proceedings Technical Report UCLA/CSD-TR-01-0023,UCLA Computer Science Department,* 2001.

[20] WR Heinzelman,A Chandrakasan,H Balakrishnan " Energy-efficient Communication Protocol for Wireless Microsensor Networks", *In proceedings of IEEE Computer Society Proceedings of the Thirty Third Hawaii International Conference on System Sciences (HICSS '00)* 2000.

[21] S Lindsey and CS Raghavendra "PEGASIS: Power-efficient Gathering in Sensor Information System" *In Proceedings of the IEEE Aerospace Conference,* Vol. 3, pp 1125-30,2002.

[22] O Younis and S Fahmy " Heed: A hybrid, Energy-efficient, Distributed Clustering Approach for Ad-hoc Networks", *IEEE Transactions on Mobile Computing,* Vol. 3(4),pp 366-69, 2004.

[23] K Akkaya and M Younis " An energy-aware QoS routing" *Proceedings of the IEEE Workshop on Mobile and Wireless Networks (MWN '03*) 2003.

Performance Analysis of Zero Cross-Correlation Code in Spectral Amplitude Coding-OCDMA

Meet Kumari[1] and Himali Sarangal[2]
[1,2]Guru Nanak Dev University,
Regional Campus, Ladhewali Road, Jalandhar, PNB
E-mail: [1]meetkumari08@yahoo.in, [2]himali.sarangal@gmail.com

***Abstract*—Optical Code Division Multiple Access (OCDMA) has bright future in communication because of its faster speed, efficiency, security and unlimited bandwidth. To reduce Multiple Access Interface (MAI) and Phase Induced Intensity Noise (PIIN) noises there are many codes in Spectral Amplitude Coding-Optical Code Division Multiple Access (SAC-OCDMA) with codes best property i.e., minimum cross-correlation property. But Zero Cross Correlation (ZCC) code used in SAC-OCDMA with zero cross correlation property is more scalable, flexible and enhance the system performance. This paper focusing on the study of ZCC code design with code most important zero cross-correlation property and analyse its performance. The result shows that ZCC code gives bit error rate of 2.7 × 10^{-20} at distance of 20km for 5 no. of users with data rate of 155Mbps. It also shows using the ZCC code leads to simplicity in system design, cost effective, reduce the MAI and PIIN noises through the zero cross correlation property. The simulation result analysed in optisystem 7.0.**

Keyword: Zero Cross Correlation (ZCC), Optical Code Division Multiple Access (OCDMA), Spectral Amplitude Coding-Optical Code Division Multiple Access (SAC-OCDMA), Multiple Access Interface (MAI), Phase Induced Intensity Noise (PIIN)

I. INTRODUCTION

The future demand of telecommunication services is moving forward day by day not to just more faster and efficient but also to high security and optical network is an efficient solution for many services due to its unlimited bandwidth capacity [1]–[3]. Optical Code Division Multiple Access (OCDMA) is a multiplexing technique of the future telecommunication as it have abilities to allow asynchronous access, high bandwidth, flexibility, cost effective and security with great advantages in local area networks and Passive Optical Networks (PON) [4], [5]. In OCDMA system one unique sequence assigned for each user is called codeword. The bit ‘1’ represent light pulse is present during that interval and bit is ‘0’ representing no light pulse. In OCDMA, Multiple Access Interface (MAI) noise causes reduction in the signal-to noise ratio of the overall system. MAI is the interference that generate due to simultaneously users’ transmission [6]. Many coding techniques are used in OCDMA to reduce MAI but spectral amplitude coding (SAC) is most effective because of assigning unique code to individual user [7]. SAC technique which is operated on bit rate, is a cost effective technique for end users [8]. Spectral Amplitude Coding-Optical Code Division Multiple Access (SAC-OCDMA) systems use cheap incoherent Light Emitting Diodes (LED) sources for SAC encoding. But also affected by Phase Induced Intensity Noise (PIIN) .PIIN noise largely limits the performance of SAC OCDMA systems. Different methods used to reduce the PIIN in SAC-OCDMA system [9], [10]. PIIN is arises due to the incoherent light mixing and incident on a photo-detector. It mainly causes because of mixing of two uncorrelated light fields with same polarization, small self-intensity noise, same spectrum and intensity [11], [12].

Low cross-correlation values reduce the MAI and PIIN effectively; hence increase the SAC-OCDMA system performance. Thus designing of code with cross-correlation is the important property with detection scheme and sender-receiver structure in OCDMA [13]. Thus ZCC code is designed to have high auto-correlation and zero cross-correlation properties. Thus it is designed with flexibility to complete the demand of multiple users with advantage of simple, efficient code design and simple system design [14].

In this paper, ZCC code design is studied. The performance of ZCC code in SAC-OCDMA is determined. The BER rate for ZCC code is determined with simulation for 5 no. of users in optisystem 7.0 at bit rate of 155 Mbps [14], [15].

II. ZERO CROSS CORRELATION CODE DESIGN

ZCC code is designed in matrix K × L means in K rows and L columns. [14], [15].

Where

K = no of users

and L = minimum code length

In ZCC code binary coefficients i.e., 0 and 1 are used.

The general form of ZCC code is given by [14], [15].

$$Z(w = i) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{1}$$

Where

[A]—represent the replication of marix w -1

[B]—represent the Matrix (w × 2w), diagonal matrix of ones with alternate of zeros matrix (w × 1) in between.

[C]–represent the (1 × w (w -1) Matrix of zero.

[D]–represent the (1 × [0 1]) matrix replication for w times.

w = weight of the matrix.

i = integer {1,2,3…∞}[15].

The basic ZCC code is designed for weight, w = 1 given as [15]

$$Z\,(w\ =\ 1)\ =\ \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{2}$$

Here the no of users are 2 and minimum code length is also 2.After transforming the code from weight = 1 to weight = 2, the matrix is given as [15]

$$Z\,(w\ =\ 2)\ =\ \begin{bmatrix}1 & 0 & 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 & 0 & 1\end{bmatrix} \tag{3}$$

Here the no. of users are 3 and minimum code length is 6.

Thus the relation between no of users and minimum code length is given by [14]

No of users = K = w + 1

Minimum code length = L = w (w + 1)

The no. of users can be increased without changing the weight by a technique called as mapping. The general form of mapping code is written as [14], [15]

$$Zm\ =\ \begin{bmatrix}Zm-1 & 0\\ 0 & Zm-1\end{bmatrix} \tag{4}$$

For weight = 1, mapping code is

$$Z2\ =\ \begin{bmatrix}Z1 & 0\\ 0 & Z1\end{bmatrix}\ =\ \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{5}$$

Where $Z1\ =\ Z\,(w\ =\ 1)\ =\ \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ from equation (2)

Here the no of users are 4 and minimum code length is 4 for weight = 1.Similarly for Z_3 „ Z_4 and so on. Thus the relationship between no of users and minimum code length for mapping technique is given by [14],[15]

No of users = Km = 2^m (K)

And code length = Lm = 2^m (L)

For incoherent light source means for LED, the broadband spectrum is divided into different wavelengths. These wavelengths divided with spacing of 0.8nm and assigned to each user [15].

TABLE I [15] WAVELENGTH ASSIGNMENT FOR ZCC CODE FOR W = 1

	λ1	λ2
Wavelength	1478.8	1479.6
User1	1	0
User2	0	1

Thus it is easier and flexible to assign wavelengths to the number of users. Adding new user (code) to the existing code wavelength will not disturb the other users [15].

III. SYSTEM PERFORMANCE ANALYSIS

A. Code Performance Analysis

Let C_m and Cn are two ZCC code sequences with weight w.

ZCC code means zero cross-correlation. So ZCC code has correlation property [14], [15].

Means

$$\sum_{i=1}^{K} Cm(i)Cn(i)\ =\ \begin{cases} w\ m\ =\ n\ autocorrelation\\ 0\ m \neq n\ crosscorrelation\end{cases} \tag{6}$$

Where K = number of user.

Autocorrelation means the comparison with same code and cross-correlation means comparison with another code. Thus for w = 2 autocorrelation is 2 in the condition of m = n and cross-correlation in 0 in condition of m ≠ n [14], [15] .

The signal to noise ratio is for SAC-OCDMA system is given as [15]

$$SNR\ =\ \frac{R^2\ P^2\frac{w^2}{L^2}}{2P\ we\ B\frac{R}{L}+4K\ TB/\ Rl} \tag{7}$$

Where R = Responstivity

P = Effective power of broadband source at receiver

w = Weight of ZCC code

L = ZCC code length

e = Electronic charge

B = Receiver electrical bandwidth

K = Boltzmann's constant

T = Absolute temperature in degrees

Rl = Receiver load resistance

Thus Bit Error Rate (BER) is given as [15]

$$BER\ =\ \frac{1}{2}erfc\sqrt{\frac{SNR}{8}} \tag{8}$$

TABLE II [14], [15] PARAMETERS USED

S. No.	Parameters	Values
1	Low Pass Bessel Filter cut off frequency	0.75 * Bit rate
2	Reference wavelength	1550 nm
3	Data bit rate	155 Mbps
4	Receiver load register	1030 Ω
5	Receiver noise temperature	300 K
6	Line–width broadband source	3.75 THz
7	Attenuation	0.2dB/km
8	Dispersion	16.75ps/nm-km
9	Dark current	10nA
10	Respostivity	1 A/W

B. Simulation Setup Analysis

The block diagram for SAC-OCDMA system implementing ZCC code is shown in Fig. 1. At the transmitter side LED is used as broadband light source.

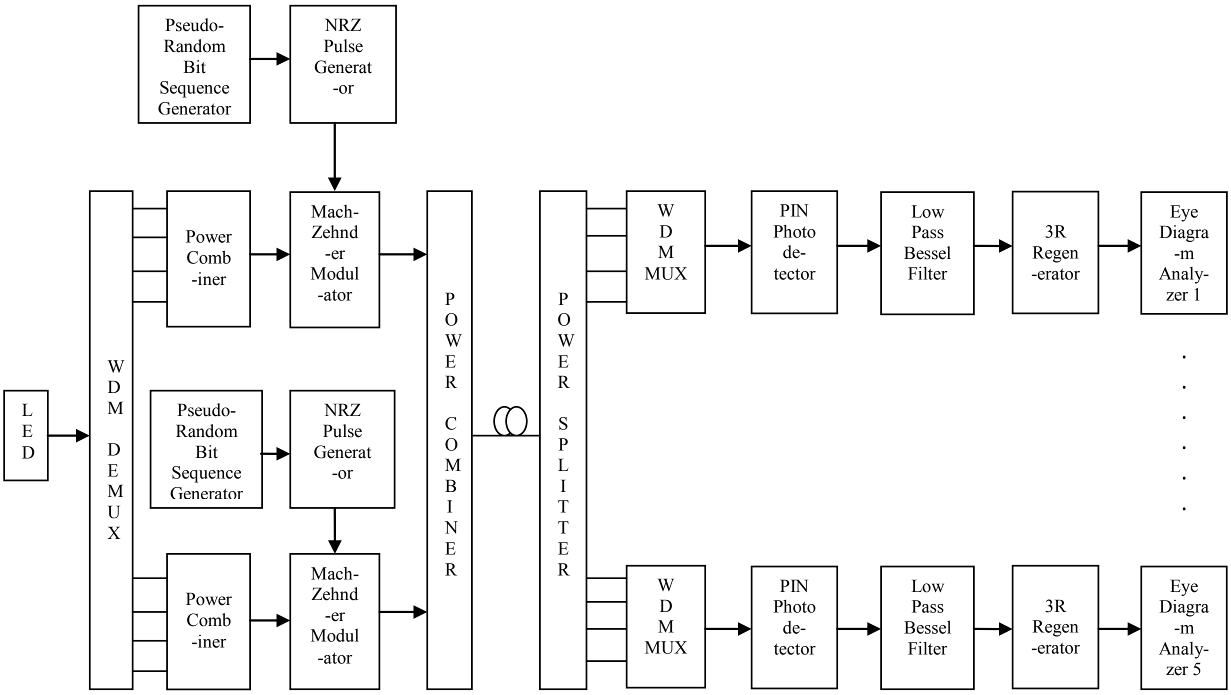


Fig. 1 Block Diagram of ZCC Code System [14]–[18]

LED broadband spectrum divided into 20 wavelengths with WDM Demultplexer with chip spacing 0.8 nm. For each code's wavelength power is combined by Power combiner. The wavelengths used in range 1478.8 nm and 1494 nm, with 0.8 nm wavelength spacing. Pseudo Random Bit Sequence (PRBS) generate data at bit rate of 155 Mbps. The NRZ signal generator used to convert logical data into electrical signal from PRBS. For modulation Mach Zehnder Modulator is used. Then it passes through optical fiber of length 20km. At the receiver side power splitter for splitting the power in WDM Mux is used. PIN is used for converting optical signal to electrical then passes through low pass filter .The large opening in eye diagram shows the immunity to noise. [14]–[17]. 3R regenerator is used for regeneration of electrical signal. The received signal finally passes through Eye diagram analyzer to take the plot of Eye pattern at the receiver end and also gives the value of Q factor, minimum BER for different number of transmitting users [18].

The wavelengths for the 4 users are shown in Table III and similarly it can be extends for 5 users.

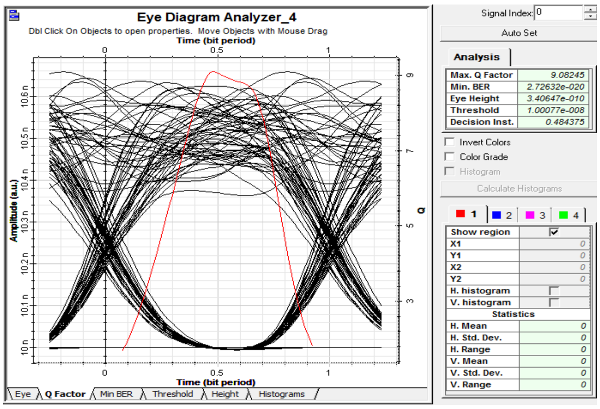
TABLE III [14]–[15] WAVELENGTH ASSIGNMENT FOR ZCC CODE FOR W = 3

	λ1	λ2	λ3	λ4	λ5	λ6	λ7	λ8	λ9	λ10	λ11	λ12
Wavelengths	1478.8	1479.6	1480.4	1481.2	1482	1482.8	1483.6	1484.4	1485.2	1486	1486.8	1487.6
User1	1	0	1	0	0	0	1	0	0	0	0	0
User2	0	1	0	0	1	0	0	0	1	0	0	0
User3	0	0	0	1	0	1	0	0	0	0	1	0
User4	0	0	0	0	0	0	0	1	0	1	0	1

IV. OBSERVATIONS

The performance of ZCC code is simulated in optisystem 7.0. In this simulation significant effects i.e., attenuation (0.25 dB/km) and the dispersion (18 ps/(nm. km) are considered .Eye diagram analyzer is used to measure both eye diagram and BER values of ZCC code in SAC-OCDMA. BER gives the error rate i.e. number of errors occurred in data link. Thus it indicate the operational performance of system. Also 10^{-9} BER is basic requirement for most light wave for error free communication system. Fig. 2 show eye diagrams taken for transmitting user 1 with single mode fiber length of 20 km at 155Mbps bit rate. It can be seen from Fig. 2 that SAC-OCDMA with ZCC code has 2.7 × 10^{-20} bit error rate. This low value of BER shows that ZCC code has better performance at wavelengths from 1478.8 nm to 1494 nm with spectral spacing 0.8 nm. Also the larger height of the eye opening at the specified time shows the immunity to noise and better system performance. Thus Eye diagram and BER values shows the better operational performance at 155mbps bit rate.

The low value of BER i.e., 2.7 × 10^{-20} gives the measurement of high quality of received signal.



Analysis	
Max. Q Factor	9.08245
Min. BER	2.72632e-020
Eye Height	3.40647e-010
Threshold	1.00077e-008
Decision Inst.	0.484375

Show region	☑
X1	0
Y1	0
X2	0
Y2	0
H. histogram	□
V. histogram	□
Statistics	
H. Mean	0
H. Std. Dev.	0
H. Range	0
V. Mean	0
V. Std. Dev.	0
V. Range	0

Fig. 2 Performance of ZCC Code in Eye Diagram Analyzer

The BER less than 10^{-9} can also be obtained with high data rate (greater than 155 Mbps) and for users greater than 5 [14]-[20].

V. Conclusion

In this paper the performance of ZCC code in SAC-OCDMA has been analysed. Code properties in code design is an important factor which enhance the performance of SAC-OCDMA system. The main property of code design is minimum cross-correlation. Thus choice of employing ZCC code increases the system performance with simple code construction and code's best property i.e., zero cross-correlation. It have several advantages like scability, flexibility, simple, can handle large no. of users with mapping method for a particular weight and immune to MAI and PIIN noises. The analysis of ZCC code proves that it is very flexible to add users in SAC-OCDMA system. The large eye opening shows its immunity to noises and BER values determined that at 155 Mbps system, high quality of received data with bit error rate of 2.7 × 10^{-20} for 20km distance is obtained.

References

[1] J.M. Nordina, S.A. Aljunida, R.A. Rahima, M.S. Anuara, A.R. Arief a, R.B. Ahmada, M.N. Saad,Performance evaluation of Fi-Wi network based on SCM–optical code division multiple access architecture, Optik124(2013) 4046–4051.

[2] Nasaruddin , Tetsuo Tsujioka, Design of strict variable-weight optical orthogonal codes for differentiated quality of service in optical CDMA networks,Computer network 52(2008) 2077–2086.

[3] Nimrat Kaur, Malti Sarangal, Effects of using RZ and NRZ modulation formats for TDM-PON system on the Transmission Characterstics for Downstream Signals,International Journal of Computer Application technology and research, volume 2, issue 6, 645–649, 2013.

[4] A.R. Arief , S.A. Aljunid, M.S. Anuar, M.N. Junita, R.B. Ahmad, Cardinality enhancement of spectral/spatial modified double weight code optical code division multi-access system by PIIN suppression, optic 124(2013), 3786-3793.

[5] A. Andueza , D. Lasaosa, D. Benito ,Analysis of electrically configurable spectral phase encoding techniques for optical CDMA,optical communication 281(2008) 5973-5981.

[6] Vishav Jyoti, R.S. Kaler,Design and performance analysis of various one dimensional codes using different data formats for OCDMA system, optik122(2011) 843-850.

[7] Thanaa Hussein Abd,S.A. Aljunid,Hilal Adnan Fadhil,Ahmad, R.A.N.M. Saad ,Design And Simulation A New Code With Zero Cross-Correlation For SAC-OCDMA Networks, Australian Journal of Basic and Applied Sciences, 6(3): 112-119, 2012, ISSN 1991-8178.

[8] Chao-Chin Yang, The application of spectral-amplitude-coding optical CDMA in passive optical networks, Optical Fiber Technology 14 (2008) 134–142.

[9] Ahmed M. Alhassana,, Nasreen Badruddina, N.M. Saada, S.A. Aljunid, A divided spectrum balanced detection technique for intensity noise reduction in SAC OCDMA systems,Optik 124(2013) 5994-5999.

[10] HamzaM. R. Al-Khafaji,S.A. Aljunid and Hilala.Fadhil, Modified-AND Subtraction Detection Technique Based on Weight Division for SAC_OCDMA System, International Journal of Computer and Electrical engg.Vol 4,No 6, December 2012.

[11] Thanaa Hussein Abd, S.A. Aljunid , Hilal Adnan Fadhil , R.A. Ahmad , N.M. Saad Development of a new code family based on SAC-OCDMA system with large cardinality for OCDMA network, optical fiber technology 17(2011) 273-280.

[12] A. Djebbari, A. Garadi, I. Dayoub, Taleb-Ahmed A ,A new code construction with zero cross correlation based on BIBD, Optik 124 (2013) 3419–3421.

[13] Hilal A. Fadhil, Syed A. Aljunid, Hassan Y. Ahmed, Hamza M.R. AlKhafaji,Variable cross-correlation code construction for spectral amplitude coding optical CDMA networks, Optik 123 (2012) 956–963.

[14] M.S. Anuar , S.A. Aljunid , N.M. Saad , S.M. Hamzah ,New design of spectral amplitude coding in OCDMA with zero cross-correlation,optics communication 282 (2009)2659-2664.

[15] Anuar Mat Safar, Syed Alwee Aljunid, Amir Razif Arief Jamil Abdullah, Junita Mohd Nordin, Mohamad Naufal Mohamad Saad, Enhancement of Zero Cross Correlation Code for Optical CDMA Network System,ICCIT 2012.

[16] Vishav Jyoti, R.S. Kaler ,Design and implementation of 2-dimensional wavelength/time codes for OCDMA, Optik 122 (2011) 851–857.

[17] Thanaa Hussein Abd,S. A. Aljunid, Hilal Adnan Fadhil,M. N. Junita and N. M. Saad, Modelling and simulation of a 1.6 Tb/s optical system based on multi-diagonal code and optical code-division multiple-access, Ukr, J,Phys,Opt 2012 ,Vol13 ,No2.

[18] Kuldeepak Singh,Dr. Manjit Singh Bhamrah, Investigation of Transmitted Power in Intersatellite Optical Wireless Communication,IRACST, ISSN:2249-9555,Vol. 2, No 3, June 2012.

[19] Irfan Ali ,Bit-Error-Rate (BER) Simulation Using MATLAB Irfan Ali, Irfan Ali / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622, www.ijera.com Vol. 3, Issue 1, January-February 2013, pp. 706-711

[20] Hilal Adnan Fadhil , S.A. Aljunid, R.B. Ahmad,Performance of random diagonal code for OCDMA systems using new spectral direct detection technique, Optical Fiber Technology 15 (2009) 283–289.

Workflow Scheduling in Mobile Cloud Computing Environment

Neha Kapoor[1] and Parveen Kakkar[2]
[1,2]Department of Computer Science & Engineering,
DAV Institute of Engineering & Technology, India
E-mail: [1]neha.kapoor2310@gmail.com, [2]parveen.daviet@gmail.com

Abstract**—Mobile Cloud Computing (MCC) is a new emerging technology that has revolutionized the way in which mobile subscribers across the globe can enjoy abundant multimedia applications on the go. MCC integrates cloud computing into the mobile environment and overcomes obstacles related to performance (e.g., battery life, storage, and bandwidth), security (e.g., reliability and privacy) and environment (e.g., scalability, availability). The scheduling of massive multimedia flows is a complicated task in the mobile cloud environment. The goal of this paper is to propose a model for job-oriented resource scheduling in a mobile cloud computing environment. The workflow is allocated or scheduled to the process which gives the available resources such as RAM, Bandwidth, MIPS etc. In this paper we construct the analysis of resource scheduling algorithms. The waiting time and turn-around time of two algorithms, viz. First Come, First Serve and Priority have been taken into consideration. From this, it has been computed that Priority Algorithm has the lowest time parameters and is the most efficient algorithm for resource scheduling.**

Keywords: Cloud Computing, Mobile Cloud Computing, Workflow, Resource Scheduling

I. Introduction

The market of mobile phones has expanded rapidly. “According to IDC [1], the premier global market intelligence firm, the worldwide Smartphone market grew 42.5% year over year in the first quarter of 2012. The growth of mobility has changed our lives fundamentally in an unprecedented way.” According to Cisco IBSG [2], close to 80 percent of the world’s population has access to the mobile phone, Android smart-phones, tablets that have brought a host of applications at the palms of people’s hands.

NIST (National Institute of Standards and Technology, USA) definition [3] from September, 2011 released in its “Special Publication 800–145” of Cloud Computing is:

“Cloud Computing is a model that enables on-demand access to a shared pool of resources (e.g., storage, applications, networks, servers and services) that can rapidly be utilized with minimal management effort or service provider interaction.”

A. *Mobile Cloud Computing*

Mobility has rapidly increased in today’s computing area. With the development of wireless technology such as WiFi and WiMAX, users are en-able to surf the internet in much easier way and that also not limited by the cables as before. Thus, the mobile devices have been accepted by more and more people as their first choice of working and entertainment in their daily lives.

“So, what exactly is Mobile computing? In Wikipedia, it is described as a form of human-computer interaction by which a computer is expected to be transported during normal usage”. [2] Mobile computing is based collectively on three major concepts: hardware, software and communication. The hardware can be considered as mobile devices, such as Android phone and laptop. The mobile computing software is comprised of the numerous mobile applications installed the devices, such as the anti-virus software, browser, g-mail and games. The communication is comprised of the infrastructure, protocols and data delivery techniques which must be transparent to the end users.

MCC is defined as “a mobile computing technology that makes use of elastic resources of various clouds and network technologies providing unrestricted functionality, storage, and mobility to mobile devices anywhere, anytime based on the pay-as-you-go principle.”

B. *Workflows*

The workflow is defined as “the automation of a business process in which tasks are passed from one participant to another to perform action, according to a set of rules.” A workflow models a process as a series of steps that simplify the execution complexity and applications management.

Scheduling is an important term when to talk about the managing requests and users in a computing environment. Scheduling is nothing but a set of task versus set of processors.

Workflow scheduling can be defined as the automation in workload scheduling where the workload is the requests and tasks generated by number of user s or clients in cloud. A scheduling can be categories into two categories: Job Scheduling and Job Mapping and Scheduling. Job Scheduling is what in which independent jobs gets scheduled among various available processors of distributed computing for optimization. A Job Mapping and Scheduling requires the allocation of multiple interacting tasks of a single parallel program in order to minimize the completion time on parallel computer system [8].

A task is an (sequential) activity that uses a set of inputs to produce a set of outputs. Processes are statically assigned to processors in a fixed set, either at compile-time or at run-time. There are two types of scheduling: static and dynamic. In static load balancing, all information is known in advance and tasks are allocated according to the prior knowledge and will not be affected by the state of the system. Dynamic load balancing mechanism tasks are allocated to the processors dynamically when they arrive. Redistribution of tasks has to take place when some processors become overloaded [7].

Job scheduling algorithms in cloud computing can be categorized into two main groups. They are Batch mode heuristic scheduling algorithms (BMHA) and online mode heuristic algorithms. In BMHA, Jobs are collected and queued into a set as they arrive and the scheduled after a fixed time slice. The main examples of BMHA based algorithms are; First Come First Served scheduling algorithm (FCFS), Round Robin scheduling algorithm (RR), Max–Min algorithm and Min–Min algorithm. In On-line mode heuristic scheduling algorithm, Jobs are scheduled as they arrive in the system. On-line mode heuristic scheduling algorithms are more appropriate for a cloud environment, as the cloud environment is a heterogeneous system and the speed of each processor varies quickly [11].

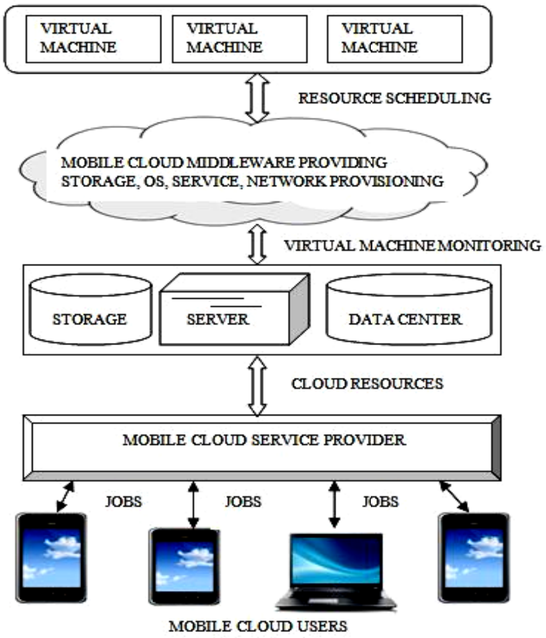


Fig. 1 Basic Resource Scheduling in Mobile Cloud Computing Environment

II. Related Works

Mayank Mishra *et al.* [2] in his paper has told that," the users of cloud services pay only for the amount of resources (a pay-as-use model) used by them. This model is quite different from earlier infrastructure models, where enterprises would invest huge amounts of money in building their own computing infrastructure. The traditional data centers results in wastage of resources during non-peak periods. The modern day data centers are shifting to the cloud to alleviate the above problem. The cloud-based data centers make resources available on demand. The cloud model provides users with a computing environment without investing a huge amount of money. As per the dynamic requirements, cloud model provides ability to dynamically scale or shrink the provisioned resources. This enables the pay as-use-go model. Thus, a cloud-based solution is an attractive provisioning alternative to exploit the computing-as-service model."

Anton Beloglazov and Rajkumar Buyya [5] have proposed "the plan for the future research work that consists of several steps listed in a table. Once the algorithms for all of the proposed optimization stages are developed, they will be combined in an overall solution and implemented as a part of a real-world cloud platform like Aneka".

Venkatesa Kumar V. and S. Palaniswami [6], in their paper, have proposed "the effective way of resource utilization and thus reduce the processing cost. Our experimental results clearly show that our proposed preemptive scheduling algorithm is effective in this regard. In this study, we present a novel Turnaround time utility scheduling approach which focuses on both the high priority and the low priority takes that arrive for scheduling."

Liang Luo *et al.* [9] in their paper have discussed "about, a new VM Load Balancing Algorithm is proposed and then implemented in Cloud Computing environment using CloudSim toolkit. He proposed an algorithm, in which the VM assigns a varying (different) amount of the available processing power to the individual application services. These VMs of different processing powers, the tasks/requests (application services) are assigned or allocated to the most powerful VM and then to the lowest. They have optimized the performance parameters such as response time and data processing time, by proposing an efficient VM Load Balancing algorithm i.e., Weighted Active Load Balancing Algorithm".

Qiang Li and Yike Guo [10] have proposed "a model for optimization of SLA-based resource schedule in cloud computing based on stochastic integer programming technique. The performance evaluation has been performed by numerical studies and simulation. The experimental result shows that the optimal solution is obtained in a reasonably short time."

W. Zhu [11], in their paper have discussed that, "by monitoring performance parameters of virtual machines in real time, the overloaded is easily detected once these parameters exceeded the threshold. Quickly finding the nearest idle node by the ant colony algorithm from the resources and starting the virtual machine can bears part of the load and meets these performance and resource requirements. This achieves the goal of load balancing and realizes the load adaptive dynamic resource scheduling in the cloud services platform."

III. Problem Formulation

"Scheduling becomes a very complicated task in a cloud computing environment in order to efficiently allocate computing resources where many alternative computers with varying capacities are available." [4] To improve the resource utilization and to meet user requirements, efficient task scheduling is required. Many users simultaneously submit lots of computing requests with different requirements to the cloud service providers. The lower and higher computing ability tasks require varying computing resources, such as cost, RAM and bandwidth. "When the cloud computing service providers receive the tasks from users, the cloud computing providers negotiate with the users on the requirements of tasks including network bandwidth, complete time, energy cost, and reliability of task. The computing resource or storage resource in a cloud computing environment can be assigned to the corresponding task according to the various scheduling policies." [5] The scheduling algorithm must be optimal and good enough to meet each user request by providing them proper resources for their execution and the task must be completed in least execution time. In our work, we have compared two scheduling algorithms to schedule the task to the various cloudlets i.e., FCFS (First Come, First Serve) and Priority Algorithm.

A. *FCFS (First Come, First Serve) Scheduling Algorithm*

FCFS follows a service policy in which the process that requests the CPU (processor) first, is allocated the CPU first without other biases or preferences.

B. *Priority Scheduling Algorithm*

Priority of jobs is an important issue in scheduling because some jobs should be serviced earlier than other those jobs can't stay for a long time in a system. So, we have proposed a prioritized Workflow Scheduling Algorithm and implemented in simulated environment. The algorithm works as follows:

Step 1: Submit the set of Jobs and set of resources say J and R respectively.
Step 2: Form the cluster of the jobs based on the certain attributes, from the given set of jobs.
Step 3: Apply the priority algorithm within each cluster to prioritize the jobs.
Step 4: Assign these clusters the computing environment which is capable of performing execution of jobs taking least time.

IV. Experimental Setup

The scheduling algorithms have been executed and resulted in a simulated environment using a simulator named CloudSim. CloudSim is a java based tool especially for the execution of cloud based application or simulation.

A. *Simulation Description*

CloudSim version 3.0.2 is used to implement Workflow Scheduling in Mobile Cloud Computing. The computer running Window 7 operating system is used for simulation.

We have implemented the "Prioritized Workflow Scheduling" model and compared it with first come first serve scheduling algorithm initially by taking into account the various parameters like waiting time and turnaround time.

B. *Virtual Machine*

A virtual machine (VM) is a software implementation of a machine (i.e., a computer) that executes programs like a physical machine [13]. In our simulation work we have considered 10 virtual machines. Table I describes the configuration of VMs.

TABLE I VIRTUAL M/C CONFIGURATION

Configuration	Value
RAM	512 MB
Image Size	10000 MB
MIPS	1000
PE's	1
Bandwidth	1000
Processor Type	Xen

C. *Cloudlet*

Cloudlet will work as Input job/task to the Cloud Environment. Cloudlet is an extension to the cloudlet. It stores all the information that is encapsulated in the Cloudlet, such as the ID of the VM running it.

V. Results and Analysis

We have done the comparison of waiting time and turnaround time of jobs when executed by applying "prioritized workflow Scheduling" versus the waiting time and turnaround time when using FCFS scheduling algorithm.

Table II shows the comparison of two algorithms based on the waiting time of jobs. Fig. 2 shows a graphical representation of comparison between two algorithms considering waiting time.

TABLE II COMPARISON WITH RESPECT TO WAITING TIME

Sr. No.	Cloudlet ID	Waiting Time Using FCFS	Waiting Time Using Priority Algorithm
1	1	0.6825	0.6528
2	2	0.4207	0.2494
3	3	1.8794	1.2507
4	4	1.4196	0.1076
5	5	0.7725	0.3817
6	500	0.1494	0.4554

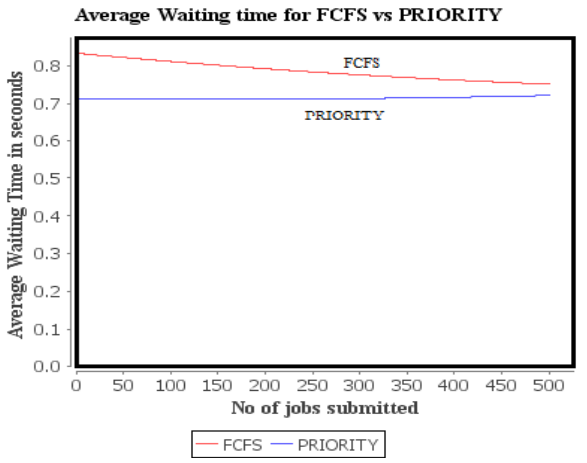


Fig. 2 Comparison of FCFS and Priority Algorithm Based on their Average Waiting Time

Table III shows the comparison of two algorithms based on the turnaround time of jobs. Fig. 3 shows a graphical representation of comparison between two algorithms considering turnaround time.

TABLE III: COMPARISON WITH RESPECT TO TURNAROUND TIME

Sr. No.	Cloudlet ID	Turn-Around Time Using FCFS	Turn-Around Time Using Priority Algorithm
1	1	28.322	13.461
2	2	13.891	13.801
3	3	14.165	13.971
4	4	14.397	14.200
5	5	29.044	13.801
6	500	0.4	0.1

VI. CONCLUSION

Scheduling is one of the most important tasks in mobile cloud computing environment. In this paper, we have analysed two scheduling algorithms i.e., FCFS and PRIORITY algorithm and tabulated various parameters.

From the observed results, it is concluded that:

1. The Prioritized Workflow Scheduling in mobile cloud computing has reduced the waiting time of jobs when compared with simple First Come First Serve algorithm.
2. The turn-around time for the jobs has been reduced in case of Priority Workflow Scheduling Algorithm as compared to FCFS algorithm.
3. The execution time has also been improved using Priority Workflow Scheduling Algorithm. Hence, Prioritized Workflow Scheduling Algorithm proves to be a more efficient Scheduling algorithm for allocating mobile cloud computing resources to the jobs submitted by the users.

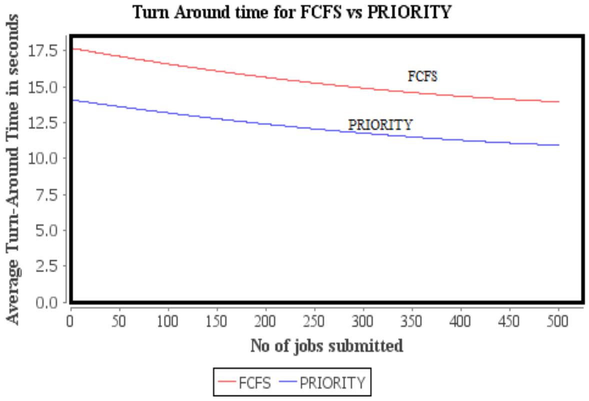


Fig. 3 Comparison of FCFS and Priority Algorithm Based on their Average Turn-Around Time

In virtual environment, we have noticed that disk space management is a critical issue. Despite of giving least execution time and cost effective, the existing scheduling algorithms do not consider reliability and availability issue. In future enhancement will propose a new algorithm considering reliability and availability issue for resource scheduling. The efficiency of the user request first may be optimized the processor and execute the request.

ACKNOWLEDGMENT

This research paper is made possible through the help and support of my guide and mentor Mr. Parveen Kakkar. I would like to thank him for his utmost support and encouragement.

REFERENCES

[1] Hitoshi Matsumoto, Yutaka Ezaki," Dynamic Resource Management in Cloud Environment", July 2011, FUJITSU science & Tech journal, Volume 47, No: 3, page no: 270-276.

[2] Mayank Mishra, Anwesha Das, Purushottam Kulkarni, and Anirudha Sahoo, "Dynamic Resource Management Using Virtual Machine Migrations", Sep 2012, 0163-6804/12, IEEE Communications Magazine, page no: 34-40.

[3] Hoang T. Dinh, Chonho Lee, Dusit Niyato, and Ping Wang. "A Survey of Mobile Cloud Computing: Architecture, Applications, and Approaches." *Wireless Communications and Mobile Computing*, vol. 13, no. 18, pp. 1587–1611, Dec 2013.

[4] Xiaocheng Liu, Chen Wang, Bing Bing Zhou, Junliang Chen, Ting Yang, Albert Y. Zomaya. "Priority-Based Consolidation of Parallel Workloads in the Cloud." *IEEE Transactions on Parallel and Distributed Systems, vol. 24, no. 9, Sep 2013.*

[5] Anton Beloglazov and Rajkumar Buyya," Energy Efficient Resource Management in Virtualized Cloud Data Centers", 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 978-0-7695-4039-9/10, IEEE, DOI 10.1109/CCGRID.2010.46, page no: 826-831.

[6] Venkatesa Kumar, V. and S. Palaniswami,” A Dynamic Resource Allocation Method for Parallel Data Processing in Cloud Computing”, 2012, Journal of Computer Science 8 (5), ISSN 1549–3636, Science Publications, page no: 780–788.

[7] Vijindra and Sudhir Shenai. A, “Survey of Scheduling Issues in Cloud Computing”, 2012, ICMOC-2012, 1877–7058, Elsevier Ltd, Doi: 10.1016/j.proeng. 2012.06.337, page no: 2881–2888.

[8] Jasmine James and Dr. Bhupendra Verma,” Efficient VM Load Balancing Algorithm for a Cloud Computing Environment “, Sep 2012, IJCSE, ISSN: 0975-3397 Vol. 4, No. 09, page no: 1658–1663.

[9] Liang Luo, Wenjun Wu, Dichen Di, Fei Zhang, Yizhou Yan, Yaokuan Mao, “A Resource Scheduling Algorithm of Cloud Computing based on Energy Efficient Optimization Methods”, 2012, 978-1-4673-2154-9/12, IEEE.

[10] Qiang Li and Yike Guo, “Optimization of Resource Scheduling in Cloud Computing”, 2010, 12th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 978-0-7695-4324-6/10, IEEE, DOI 10.1109/SYNASC.2010.8, page no: 315–320.

[11] W. Zhu *et al*. “Multimedia Cloud Computing.” *IEEE Signal Processing Magazine*., vol. 28, no 3, pp. 59–69, May 2011.

Wireless Sensor Network: Threat Models and Security Issues

Reenkamal Kaur Gill[1], Priya Chawla[2] and Monika Sachdeva[3]
[1,2,3]Department of Computer Science and Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab, India
E-mail: [1]reenkamalgill@gmail.com, [2]piyachawla12@gmail.com, [3]monika.sal@rediffmail.com

***Abstract*—Wireless Sensor networks is new emerging technologies that involves the deployments of tiny devices which are equipped with sensors communicate with each other over a specific geographical area to provide a collaborative measurement. This sensor arrangement can be used for specific purposes such as smart cities, smart agriculture, etc. Wireless Sensor Networks are prone to various kinds of threats and attacks. In this paper, we analyze different threat models, security issues and attacks that should be resolved to make the sensor network secure and smooth going.**

Keywords: Sensor, Wireless Sensor Networks, Attacks, Threat Models, Security

I. Introduction

Wireless Sensor Network is composed of large number of sensor nodes and the basic idea of sensor network is to deploy the sensor nodes in some geographical area which are capable of monitoring and recording the physical conditions of environment like temperature, sound, pollution level, humidity, etc. and for several other purposes like target tracking, surveillance, etc.

Unstructured WSN: It is a network that contains a large number of sensor Wireless Sensor Network is categorized as: an Unstructured and Structured Wireless Sensor Networks.

Nodes and they can be automatically organized to form an ad-hoc network.

Structured WSN: It has a pre-planned criteria that how to deploy the sensor nodes in large geographical area.

Hence, on the whole we can say that Structured WSN has an advantage over Unstructured WSN that it has lower management and maintenance cost.

Various features of WSN that attracts researchers to pay attention towards various issues related to these networks. But if we analyze previous researches, we could observe that routing strategies of WSN have been given much more priority. But in this paper, we will discuss about the security issues of WSN as well as their challenges. In the second section we will discuss about the various elements of WSN and in the next section various threat models are discussed. In the third section we will emphasize on various dimensions of security like confidentiality, integrity, authentication and data freshness. In the last section various attacks on routing protocols will be presented.

II. Elements of WSN

Typical elements of wireless sensor network are:

Node: It is an autonomous device equipped with sensors. Node includes a radio transceiver along with an antenna, a microcontroller, an interfacing electronic circuit, and an energy source usually a battery. e.g., WaspMote.

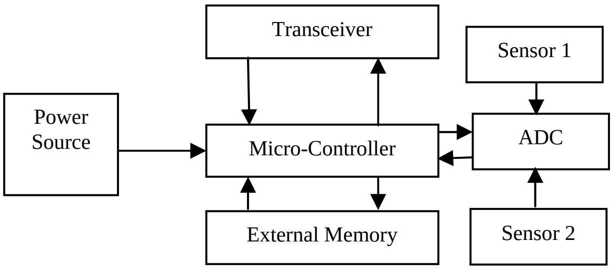


Fig. 1 Typical Architecture of the Sensor Node

Data Gatherer: This is a data capture device and it should be connected to external system in order to transmit sensor value.

e.g., Mashlium Xtreme.

External System: It is a data storing and managing centres. In case we want to store some data, we can use a non-volatile memory with some available space (e.g., EEPROM with 3 KB available) or maybe we can use the SD card (2 GB available) to store all the sensor values.

III. Threat Models

Attacks on Wireless Sensor Network can be categorized into various categories on the basis of certain criteria. In the first category attack can be: Mote Class or Laptop Class [1].

In mote class attacker can interact with only few sensor nodes where the entire sensor nodes must have similar capabilities, whereas in case of a Laptop Class attacker can interact with more powerful devices like PDA’s, Laptops etc.

A Laptop Class adversary can produce a huge amount of damage than Mote Class. Mote class adversary can effect only within small geographical area, but on the other hand Laptop Class adversary could have an effect on the entire network and even could block the entire sensor network.

Another classification of attack on Wireless Sensor Network can be: Insider or Outsider Attack.

In case of Insider attack the attacker has access to that node which has all the secret keys and is capable of participating in all the communications.

In Outsider attack attacker has no access to Wireless Sensor Network. It is done by the unauthorized node that eavesdrop the packets exchanged between the sensor nodes during their communication.

Next classification of attacks is based on Network Layer which are: Attacks at Physical Layer, at Data Link Layer and at Network Layer.

At Physical Layer attacker mainly exhaust the resources available by transmitting the radio signals on a Wireless Channel.

At Data Link Layer the attacker violate the predefined protocols of the Link Layer. This kind of attack also leads to Denial of Service attack.

At Network Layer attacker threatens the sensor applications and services. In this Localization and Aggregation are used to prevent from this attack.

TABLE 1: WSN THREATS IN LAYERS [2]

Layers	Attacks
Physical	1. Denial of Service (DoS) 2. Tempering
Data Link	1. Jamming 2. Collision
Network	1. Sybil Attack 2. Wormhole Attack 3. Sinkhole Attack 4. Flooding
Application	1. Desynchronization 2. Aggregation based attacks

IV. SECURITY REQUIREMENTS IN WSN

The different security concerns of Wireless Sensor Network are as follows:

1. *Data Confidentiality:* It means the content of the message when transmitted across the network must remain confidential i.e. only the intended receiver and no one else should be able to read the message. Hence encryption is used for effective and secure communication in which data is encrypted into secret words.
2. *Data Integrity:* It means data must reach the destination without being changed by the adversaries or Attackers. Data Integrity ensures that the data has not been changed during the transmission, neither accidentally or intentionally. Checksum is used for data integrity.
3. *Data Authentication:* It is the fundamental requirement for security in WSN. Attacks in the sensor networks do not just involve the alteration of packets; adversaries can also inject additional false packets [3]. In message authentication, receiver needs to be sure of the sender's identity as an adversary can change the entire data. So the receiver needs to be assured that whatever data used in Decision making process comes from an authorized source or not.
4. *Data Freshness:* Data freshness [4] ensures that data should be recent and no old messages have been replayed. This requirement is essential when shared key strategies are used. So there is a great need to get renew the shared keys time to time. As it takes a little bit time to propagate the shared keys over the entire network during that time adversary can perform a replay attack. To tackle the problem of the replay attack timestamp is added to the message.

V. MAJOR ATTACKS IN WSN

As most of the routing protocols for WSN are very simple, so they are more vulnerable to attacks [5]. Attacks on Network Layer protocol fall into one of the following categories:

1. *Denial of Service (DoS):* Denial of Service (DoS) [6] attack is produced by malicious nodes or users. The main intention behind this attack is that it is an attempt to make network resources unavailable to its legitimate users. As the network is flooded with huge requests by an adversary so that legitimate user cannot access the services of the network. In wireless sensor network several types of DoS attacks might be performed at different layers. At Physical Layer, DoS attack can be known as Jamming. In this, adversaries continuously transmit radio signals as well as high energy signals so that wireless medium could be blocked. Jamming is further of two types: Constant Jamming and Intermittent Jamming. In case of Constant Jamming, there is a complete jamming of the entire network, whereas in case of intermittent jamming, sensor nodes are capable of communicating data periodically but not consistently.

TABLE 2: DENIAL OF SERVICE ATTACKS AND DEFENSES TO COMBAT AT DIFFERENT PROTOCOL LAYERS [7]

Protocol Layer	Attacks	Defenses
Physical	Jamming	Sleep
	Node Destruction	Hide nodes or tamper proof packaging
MAC (Medium Access Control)	Denial of Sleep	Sleep, authentication and anti-replay
Network	Spoofing, replaying	Authentication, anti-replay
	Hello floods	Geographic routing
	Homing	Header Encryption
Transport	SYN flood	SYN cookies
	De synchronization attack	Packet authentication
Application	Path based DoS	Authentication and anti-replay protection
	Reprogramming attacks	

1. *Selective Forwarding:* It is also a network Layer attack. In this, an adversary usually forwards some of the packets and drops rest of the packets containing vital information. This degrades the quality of service in WSN. If somehow attacker discards all the packets, then the receiver node becomes conscious that there must be some obstacle in between, so neighboring nodes decide to take another route. But to overcome this doubt adversary forwards a selective packets to the node rather than dropping all the packets.
2. *Blackhole/Sinkhole Attack:* In this attack the major intention of the malicious node (Blackhole [8]) is to attract the maximum traffic towards itself. An adversary makes assures to all the sensor nodes that it is also a compromised node and it will provide them the best quality route and even the shortest path to the base station. Then all the neighboring nodes of the adversary will start transmitting the packets to the adversary and when the whole of the traffic reaches the adversary, it can do anything with that information. Even it can perform selective forwarding attack i.e., to drop the crucial data and forwards the rest of the irrelevant packets to the base station.
3. Sensor networks are much more prone to sinkhole attacks as they have the common destination and a compromised node needs only to assure that it will provide high quality routes to the base station just to attract the maximum traffic towards itself.
4. *Hello Attack:* In this attack, malicious node having high radio transmission range broadcasts HELLO message to the neighboring sensor nodes to make them assure that it is also a legitimate node as well as it will provide shortest route to the base station. As a result, while sending the packets to the base station sensors nodes packets pass through the malicious node because it has made them an illusion that it is their legitimate neighboring node and hence get all the relevant data and attack the sensor network.
5. *Sybil Attack:* This type of attack mainly occurs in peer to peer network and detection of it is very difficult. We define Sybil attack as a malicious node which forges the false identity of many legitimate nodes [9], [10]. Whenever there is a communication between the two legitimate nodes adversary node occurs in between the interaction and proves the sender node that it is the one that the sender wants to exchange the data with by using the identity of receiver node. Adversary node takes all the information and hence can use selective forwarding, degrading of the data, etc. Newsome *et al.* [10] used radio resource testing to detect the presence of Sybil node in the sensor network and showed the probability to detect the existence of Sybil attack.
6. *Wormhole Attack:* Wormhole [11] is a critical attack in which attacker connects two distant points in the network using a low latency communication link called wormhole link [12]. Once the link is established the adversary records the packet at one location in the network and replays them at the other end. This type of attack is a significant threat to the sensor network as it could even be performed at the initial phase when sensors discover their neighboring information.

VI. Conclusion

Wireless Sensor Networks would be widely deployed in future mission critical applications. So security related issues in wireless sensor networks have become an important part of research in present scenario. In this paper we have described various security requirements in wireless sensor networks and also emphasized on various attacks related to wireless sensor networks.

As most of the attacks against security in this network are caused by the insertion of false information by the adversary node, so there is a great need of detecting the false reports and to develop such a mechanism that detects this is a great research challenge.

References

[1] Chris Karlof and David Wagner, “Secure Routing in Wireless Sensor Networks: Attacks and Countermeasures,” University of California, Berkeley.

[2] Abhishek Jain, Kamal Kant and M.R. Tripathy, “Security Solutions for Wireless Sensor Networks,” Amity University, India, In Second International Conference on Advanced Computing and Communication Technologies.

[3] Y. Wang, G. Attebury, and B. Ramamurthy, “A Survey of Security Issues in Wireless Sensor Networks,” *IEEE Commun. Surveys Tutorials*, vol. 8, pp. 2–23, year 2006.

[4] Ian F. Akykildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci,“A Survey on Sensor Networks”, IEEE Communication Magazine, year 2002.

[5] Al-Sakib Khan Pathan, Hyung-Woo Lee, and Choong Seon Hong, “Security in Wireless Sensor Network: Issues and Challenges,” *Kyung Hee University Korea,* Feb. 2006.

[6] Blackert, W.J., Gregg, D.M., Castner, A.K., Kyle, E.M., Hom, R.L., and Jokerst, R.M., “Analyzing interaction between distributed denial of service attacks and mitigation technologies”, Proc. DARPA Information Survivability Conference and Exposition, Volume 1, 22-24 April, 2003, pp. 26 – 36.

[7] Doddapaneni, Krishna Chaitanya and Ghosh Arindam “Analysis of DoS attack on WSN using Simulation.”

[8] Culpepper, B.J. and Tseng, H.C., “Sinkhole intrusion indicators in DSR MANETs”, Proc. First International Conference on Broad band Networks, 2004, pp. 681 – 688.

[9] Douceur, J. “The Sybil Attack”, 1st International Workshop on Peer-to-Peer Systems (2002).

[10] Newsome, J., Shi, E., Song, D, and Perrig, A, “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.

[11] Hu, Y.-C., Perrig, A., and Johnson, D.B., “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.

[12] Ritesh Maheshwari, Jie Gao, Samir R Das, “Detecting Wormhole Attacks in Wireless Sensor Networks using Connectivity Information,” In IEEE INFOCOM 2007, Alaska.

[13] Mohammad Sadeghi, Farshad Khosravi, Kayvan Atefi and Mehdi Barati, “Security Analysis of Routing Protocols in WSN,” Jan. 2012.

A Survey of VANET Routing Protocols

Aarja Kaur[1] and Sabia[2]
[1,2]*Department of Computer Sc. & Engineering,*
Guru Nanak Dev University, Regional Campus, Jalandhar, India
E-mail: [1]aarja.kaur@yahoo.com, [2]sabiajal@gmail.com

***Abstract*—Vehicular Ad-hoc network (VANET) is an on-demand wireless network that provides for communication between moving vehicles (V2V) and between vehicle and infrastructure (V2I). VANET implements intelligent transportation system (ITS) and aims to optimize traffic flow, improve road safety and reduce congestion. The communication depends on routing. The sporadic connectivity and sudden changes in network topology are the characteristics of VANET that make routing a challenging task. This paper gives a brief overview of routing protocols in VANET, their issues which are under research.**

Keywords: VANET, Routing, WAVE, ITS

I. Introduction

VANET is a subclass of mobile ad-hoc networks (MANETs). It is a self organizing network without any physical infrastructure. VANET allows the fast moving vehicles to exchange real-time information that can assist the drivers to avoid any situation like-accidents, traffic jams, etc .With the rapid increase in the vehicular traffic on roads, the corresponding increase in accidents created a security issue that drew the attention of researchers towards VANET. Dedicated short range communication (DSRC) facilitates the wireless communication in VANET. DSRC is IEEE 802.11p standard and is a MAC protocol operating at 5.9 GHz [13]. IEEE has standardised the whole communication stack that is referred to as wireless access in vehicular environments (WAVE). VANET provides a wide variety of applications for both safety and non-safety purposes. The major application of VANET is ITS [14]. In addition several value added services such as enhanced navigation; automated toll payment, internet access, and location based services are also provided. In VANET, each vehicle is equipped with devices that allow it to send, receive and exchange information with other vehicles or road side units.[8] Facilitating communication among the vehicles and developing an efficient routing protocol in VANET is a challenging task due to the following reasons: signal fading due to the presence of obstacles (buildings etc.), bandwidth constraints, high mobility of the vehicles and the speed depends on the traffic signs and signals. High mobility results in frequent fragmentation in the network [11]. The routing protocols devised for use in VANET can be categorised under topology based and position based routing.

The rest of the paper is organised as follows. In section II, a brief introduction to the different VANET protocols is given. Section III provides the related work in routing. Section IV provides a survey on recent routing protocols and various issues. Section V concludes the paper.

II. Routing Protocols

VANET's involve vehicles that act as both mobile nodes and routers for the purpose of data dissemination and enable ITS. Routing is a major research challenge in VANET because of high mobility and abrupt changes in topology. Research is being done for designing an efficient routing protocol. Due to the similarities between Mobile Ad-Hoc networks (MANET) and VANET, the traditional ad-hoc routing protocols for MANET are also applied to VANET. These include the topology based routing protocols [14]. Further other routing protocols devised for VANET fall under the following categories-Position based, cluster based, broadcast and geocast routing. We briefly describe topology based and position based routing protocols.

A. *Topology Based Routing*

These protocols discover the route based on the link information and maintain it in a table. They are further branched into three categories namely, proactive, reactive and hybrid protocols [10].

B. *Proactive Protocols*

Table driven routing protocol is another name for proactive protocols. Any change in the topology of the network is recorded by the nodes in their respective tables and the tables are periodically exchanged with the neighbours. Although these protocols consume much bandwidth for periodic updates of topology but delay involved for initial route discovery is almost negligible. Routing protocols that fall under this category are:

1. Destination sequence distance vector routing (DSDV).
2. Optimized link state routing (OLSR).
3. Source-tree adaptive routing (STAR).
4. Fisheye state routing (FSR).
5. Reactive protocols.

These protocols are known as on-demand routing protocols since they modify the routing table periodically only in case there is some data to send. Flooding process is utilized by these protocols for the purpose of route discovery which causes routing overhead. Some of the protocols under this category are:

1. Ad-hoc on demand distance vector routing (AODV).
2. Dynamic source routing (DSR).

3. Temporally-ordered routing algorithm (TORA).
4. Hybrid protocol.

The overhead in proactive routing and the initial route discovery delay in reactive protocols led to the discovery of hybrid protocols. In this protocol, reliability for route discovery and maintenance is provided by dividing the network into zones [15]. The protocol under this category is:

1. Zone routing protocol(ZRP)

C. Position Based Protocols

These protocols assume that each node has knowledge about its physical/geographic position by GPS. The physical location is used to select the next forwarding hop and hence no global route between the source and the destination needs to be maintained. [9]Some of the protocols that fall under this category are:

1. Greedy Perimeter Stateless Routing (GPSR).
2. Geographical Source Routing (GSR).
3. Anchor-based street and traffic aware Routing (A-STAR).

III. Literature Survey

Considering the growing need of VANET for security purposes, various researchers have proposed different routing protocols for efficient data delivery with minimum time delay and performed comparisons to find the best among the already existing protocols.

1. K. Prasanth, Dr. K. Duraiswamy [1] in 2009 come up with a greedy position based routing approach. In this approach, the source node identifies its neighbour nodes in its transmission range and that are moving in the direction of the destination node. Finally, the specific edge node within the limited transmission range is chosen as the next hop. Results have shown that this proposed edge node based greedy routing out performs the GPSR and PDGR in terms of packet delivery ratio.
2. Jerome Haerri [2] in 2009 used Vehicle Mobility Model to study the characteristics for the purpose of evaluating the working of AODV and OLSR for VANET in city environment. The results showed that OLSR had better performance over AODV in city scenario.
3. Shaikhul Islam Chowdhury,Won-lee, Youn-Sang Choi [3] in 2011 used different mobility models to evaluate the performance of reactive routing protocol like DSR, AOMDV, AODV by considering different performance metrics, the simulation showed that AOMDV performs better than AODV and DSR in terms of end to end delay.
4. Dharmendra Sutariya, Dr Shrikant Pradhan [4] in 2012, proposed a new version of AODV called improved AODV (IAODV) that ensured timely and accurate delivery of information to the vehicle drivers. Simulation results showed that TAODV outperforms AODV in city scenarios in terms of end-end delay, packet loss ratio, and packet delivery ratio.
5. Jamal Toutour, Jose Garcia Neito [5] in 2012 proposed on intelligent OLSR routing protocol optimization for VANET. They presented a solution to the optimization problem in order to tune the OLSR used in MANET to fit the characteristics of VANET. The quality to service of OLSR significantly improved by changing the configuration parameters.
6. Mohammad Al-Rabayah, Robert Malaney [6] in 2012 proposed a new scalable hybrid routing protocol in order to overcome the excessive overhead resulting from link failures due to high mobility of vehicles. HLAR succeeds in reducing the routing overhead compared to the standard reactive and geographic routing protocol.
7. Qinlin, Changle li, Xin Wang [7] in 2013 realized the routing issues that arise when the known VANET protocols like proactive, reactive, position based are applied in 3D scenarios of VANET. These protocols are mainly analyzed and designed based on ideal plane scenarios so when applied to 3D scenarios, server problems occur. A 3D scenario oriented routing protocol is devised for the 3D scenarios like tunnel, ramp.TDR is tested to be better than GPSR in terms of average hops, end-end delay and delivery ratio.
8. Baber Aslam, Faisal Amjad [12] in 2013 came up with a Privacy-enhancing Multilayer Trajectory based Routing Protocol (PMTR) for VANET. Setting up VANET when the number of VANET-enabled vehicles and road side units are limited, is a complex task. The routing protocols available will have poor performance since the connection among the vehicles (V2V) will be frequently disrupted. Initial deployment of VANET is of concern in the near future. PMTR routes messages using the past traffic history and trajectory information provided by vehicles. This protocol uses carry and forward paradigm and previous traffic statistics to preserve privacy. It is found that PMTR has less overhead and provides better privacy as compared to the other geographical protocols.

IV. Issues

TABLE 1 PROS AND CONS OF VARIOUS ROUTING PROTOCOLS

Protocol	Pros	Cons
Proactive Protocols	1) Initial route discovery is not required. 2) Low delay in real time applications	1) Bandwidth is wasted in storing unused paths 2) Significant overhead in periodically sharing tables.
DSDV	1) Loop free path to the destination is achieved	1) Full dump packets decrease the bandwidth utilisation as only updates are sent not the complete information. 2) Incremental packets increase the overhead as send frequently.
OLSR	1) Well suited for high density networks	1) Requires a routing table for all possible routes, leading to overhead and constrains scalability.
Star	1) Reduces overhead as no frequent updates are required.	1) Requires large memory for maintaining large trees for the network.
Reactive Protocols	1) Beaconless so it saves the bandwidth. 2) Less overhead by maintaining only the currently active routes.	1) Delay in route discovery and maintenance. 2) Not suitable for large scale networks.
AODV	1) The path to the destination is updated using the destination sequence number. 2) Low memory requirements and route redundancy. 3) Reduces flooding and network overhead. 4) AODV responds to the link failure in the network. 5) Applicable to large scale ad hoc network.	1) Connection setup and establishment of route is time consuming. 2) Extra bandwidth is needed for periodically sending beacon messages.
DSR	1) No periodic update required in DSR. 2) A node can save more than one route to a destination .The cache roué can be used in case a route breaks.	1) In large networks, byte overhead results from the huge amount of route information stored. 2) Performance worsens with increasing mobility. 3) Broken links cannot be repaired locally.
Hybrid Protocols-ZRP	1) Reduces network overhead caused by proactive protocols and handles delay caused by reactive protocols	1) Not suitable for VANET where there is dynamic change in topology.
Geographical Protocols	1) Route maintenance is not required. 2) Stores information of source, destination & neighbouring nodes.	1) Availability of position determining services is a must. 2) Due to the absence of satellite signal, GPS device is unable to function.
GPSR	1) Greedy and perimeter forwarding provide better routing decisions. 2) Forwarding packet decisions are made dynamically. 3) Robust in highly dynamic network.	1) High mobility can make a node unable to maintain information of its next hop neighbour. 2) Beacons may be lost due to bad signal.
GSR	1) GSR has a better packet delivery ratio as compared to AODV & DSR. 2) GSR is scalable than AODV & DSR.	1) The situation such as sparse network where there are not enough nodes for forwarding packets is ignored by this protocol.
A-STAR	1) A new local recovery strategy that is more practical/applicable in city environment is used by A-STAR.	1) A-STAR has a lower Packet delivery ratio as compared to GSR & GPSR.

V. Conclusion

The main goal of this survey was to study the routing protocols proposed for VANET. The article provides a review of several traditional routing protocols devised for use in VANET, including reactive and proactive protocols. The issues in these protocols are summarized in Table 1. The proactive and reactive protocols have inherent supposition of network connectivity. Disconnection and discovery of new nodes is ignored hence these protocols are not well suited for VANET. Several improved versions of the traditional routing protocols like AODV, OLSR, hybrid protocols are proposed for improving their performance, details of which are in [4, 5, and 6]. Research is being carried out for designing protocols and algorithms that can perfectly fit in the characteristics of VANET. The main limitation of many protocols is long time delay and the number of retransmissions. It has been found that position-based routing, geocasting are most promising for data dissemination in VANET as given in [13]. The survey shows that a routing protocol works well only in a particular scenario like city, urban environment etc. There is no universal protocol which is suitable for all VANET's application scenario. A specific routing protocol is needed to satisfy the requirements of a particular VANET application, which is a difficult task.

In future, work needs to be done to generalize algorithms to fit in different scenarios. Instead of designing new protocols, a protocol should be able to adapt to the abrupt changes in network and diverse mobility patterns. Work can also be done for making routing more secure as privacy is a major issue in VANET's.

REFERENCES

[1] K. Prasanth, Dr. K. Duraiswamy IEEE Senior member, K.Jayasudha and Dr.C.Chandrasekar," Minimizing End-to-End Delay in Vehicular Ad Hoc Network using Edge Node Based Greedy Routing", IEEE 2009.

[2] Jerome Haerri, Fethi Filali, Christian Bonnet, "Performance Comparison of AODV and OLSR in VANETs Urban Environments under Realistic Mobility Patterns" 2009.

[3] Shaikhul Islam Chowdhury, Won-Il Lee, Youn-Sang Choi, Guen-Young Kee, and Jae-Young Pyun," Performance Evaluation of Reactive Routing Protocols in VANET " Asia-Pacific Conference on Communications (APCC) 2011.

[4] Dharmendra Sutariya, Dr. Shrikant Pradhan," An Improved AODV Routing Protocol for VANETs in City Scenarios" IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM-2012) March 30, 31, 2012

[5] Jamal Toutouh, José García-Nieto, and Enrique Alba," Intelligent OLSR Routing Protocol Optimization for VANETs" IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 61, NO. 4, MAY 2012

[6] Mohammad Al-Rabayah, Robert Malaney," A New Scalable Hybrid Routing Protocol for VANETs" IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 61, NO. 6, JULY 2012

[7] Qinlin, Changle li, Xin Wang, Lina Zhu," A Three-dimensional Scenario Oriented Routing Protocol in Vehicular Ad Hoc Networks", IEEE 2013.

[8] Saif Al-Sultan , Moath M. Al-Doori, Ali H. Al-Bayatti, Hussien Zedan," A comprehensive survey on vehicular Ad Hoc network", Journal of Network and Computer Applications,21 February 2013.

[9] Pooja Rani , Nitin Sharma, Pariniyojit Kumar," Performance Comparison of VANET Routing Protocols ", IEEE 2011.

[10] Jagadeesh Kakarla, S Siva Sathya1, B Govinda Laxmi, Ramesh Babu B," A Survey on Routing Protocols and its Issues in VANET", International Journal of Computer Applications, August 2011.

[11] Amandeep Kaur, Issues and Challenges of Routing Protocols in VANET-Survey, National Conference on Innovations in Engineering and Technology (IEIT-2014), March 2014.

[12] Baber Aslam, Faisal Amjad ,Cliff C.Zou,"PMTR:Privacy-enhancing Multilayer Trajectory-based Routing Protocol for VANETs", IEEE Military Communications Conference 2013.

[13] Yousef-Awwad Daraghmi, Chih-Wei Yi,"Forwarding Methods in Data Dissemination and Routing Protocols for Vehicular Ad Hoc Networks",IEEE Network 2013.

[14] Fan Li , Yu Wang,"Routing in Vehicular Ad Hoc Networks: A Survey", IEEE Vehicular Technology Magazine,2007.

[15] Marwa Altayeb1 and Imad Mahgoub2," A Survey of Vehicular Ad hoc Networks Routing Protocols", International Journal of Innovation and Applied Studies ISSN 2028-9324 Vol. 3 No. 3 July 2013, pp. 829-846.

A Survey on Mobility Management Techiques in Vehicular Ad-hoc Network

Richa Sharma[1] and Jyoteesh Malhotra[2]
[1,2]*Department of Computer Sc. & Engineering*
Guru Nanak Dev University Regional Campus, Jalandhar, India
E-mail: [1]rixs11111991@gmail.com, [2]jyoteesh@gmail.com

***Abstract*—It came to the light of the world that last few years have brought a boom in the wireless communication. Continuous progresses in this field have opened new research field in computer networking aimed at enhancing data network connectivity to environment where the wired networks are not that practical. Due to importance of related applications vehicular networks are attracting most of the researchers and its mobility management is a challenging task due to small coverage range of 802.11 Access Point (AP) and fast mobility of the vehicles. Still now these traditional mobility management techniques cannot meet the requirement of vehicular networks and performance degrades. This article presents a comprehensive survey on mobility management solutions in vehicle to infrastructure (V2I) and vehicle to vehicle (V2V) communication protocols in vehicular ad hoc network (VANET). Analysis is done on the critical issues over which the future vehicular applications should be deployed. The existing mobility management schemes are also reviewed. At the last, several open issues in mobility management for vehicular networks are outlined.**

Keywords: V2I Communication, V2V Communication, Handoff and Location Management, Research Issues

I. INTRODUCTION

VANET is an intelligent vehicular Ad-hoc network which uses WiMAX IEEE 802.16 and WiFi IEEE 802.11for efficient communication between vehicles with varying mobility. VANET is a type of Mobile ad-hoc network which provides communication among vehicles and vehicles and fixed equipments nearby usually these are called as roadside equipments. The key difference between MANET and VANET is as VANET as special mobility pattern and rapidly changing topology. [2] We cannot implement existing routing protocols of MANET in VANET. VANET are widely used to support the growing number of wireless products which can be used in vehicles. VANET is a special type of mobile ad-hoc network which is divided into V2I and V2V networks. To introduce this many researchers has introduced Media Access Control protocols to improve VANET working. Evaluated from cellular networks, mobility management in VANET has seems to be a tough issue to support seamless communication. Mobility management has two types of management, one is handoff management and other is location management. Handoff management helps to regulate the active connections between mobile nodes when they change its point of attachment. Location management functions to continuously update and track the current location of mobile nodes. Different types of wireless standards used: 802.11 (WLAN), 802.13 (WiMax), 802.15 1 (Bluetooth), 802.15 4 (Gigbee). [1] Due to various architectural differences between V2I and V2V communications, their mobility management patterns are designed differently. V2I communication is designed based on the internet mobility management protocols due to interoperability and compatibility reasons (e.g., mobile IPv6). [3] [7] For mobility management in V2V communication, it mainly emphasis on route discovery, maintenance and recovery. The rest of paper is organized as follows. Next section gives the brief introduction of the mobility management in VANET. In section III an exhaustive related work has been presented for various different mobility management techniques. Section IV provides the open issues and future scope before the paper is finally concluded in section V.

II. MOBILITY MANAGEMENT IN VEHICULAR NETWORKS

In this section ITS (intelligent transport system) of VANET is described and V2I and V2V communication scenarios are discussed.

A. ITS

VANET is a very critical part of ITS i.e., intelligent transport system. ITS provides the user various innovative and resourceful services for different modes of transportation. V2V and V2I communication in VANET is implemented on ITS. Various systems of ITS:

1. Vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communication.
2. Advanced Transportation Management System (ATMS). e.g., traffic management centre.
3. Advanced Public System Transportation (APTS). e.g., electronic fare payment.
4. Advanced Traveler Information System (ATIS) e.g., parking information.

B. Mobility Management for V2V Communication

In this section discussion is done on the various problems arise in different mobility management for V2V communication. Mobility in VANET is managed through route discovery, maintenance and its recovery. [2]

1) Topology management

Two types of schemes belong to this topology management. In proactive scheme its sends signaling messages periodically to know the scenario of which type of topology. In reactive scheme it only on demand obtain the topology information. As the main problem which arises in VANET is its large topology. This problem is resolved by using cluster based topology control in which vehicles are grouped together into multiple clusters. Also for interconnection communication COIN network was proposed [9]. Also a prediction based topology was proposed [4].

2) *Location management*

Due to large latency and overhead ad hoc routing protocols are unable to apply on VANET. Various geographical routing protocols are used to solve this issue like greedy perimeter stateless routing (GPSR), geographical routing algorithm (GRA). Further flooding based approach and rendezvous based approach is used in location management in VANET which gives rise to region based location service management protocol (RLSMP) which support both scalability and locality awareness [10]. In 2002 mobile internet protocol version 4 (MIPv4) was introduced to achieve seamless handover because of problems like IP addresses, weak security, routing problems in 2004 mobile internet protocol version 6 (MIPv6). It removes all the drawbacks of MIPv4. Either for further efficiency improvement hierarchical mobile internet protocol (HMIv6) was produced in 2005 [7]. This HMIv6 has produced a new concept known as Mobility Anchor Point (MAP) which easily manages user location. MAP provides macro mobility and micro mobility management techniques for location management. Also HAWAII [8], Proxy MIPv6 (PMIPv6) [10], Cellular IP [7] was further introduced.

3) *Handoff management*

Handoff management is needed for successful communication of vehicles. So in VANET special rerouting is done to construct a new path from sender to destination. As packet loss and packet delay are the two basic problems in handoff. Somehow this problem is solved by introducing WIMAX Mobile Multi hop Relay MMR [5]. This technique provides good communication even when vehicles are on high speed freeways.

C. *Mobility Management for V2I Communication*

1) *Host mobility management*

Researchers until have given many solutions for host specific mobility in VANET. As in link layer when mobile node moves specifically between the accesses points in a common subnet, this mobility are handled by link layer protocols e.g., WLAN. Also Fast DLMAP-IE was introduced to reduce downlink traffic reception during handoff process. Some research work has been done on the various host specific standard mobility and handover management schemes in VANET listed below in tabular form.

TABLE I COMPARISON OF VARIOUS MOBILITY MANAGEMENT TECHNIQUES

	MIPv6	**FMIPv6**	**FHMIPv6**	**IFHMIPv6**	**SCTP**	**SIP**
Route Optimization Support	Yes	Yes	No	No	Yes	Yes
Protocol Layer	L3	L3	L3	L3	L4	L4
Cross Layer Information Required	No	Yes	Yes	Yes	No	No
Handover Latency	High	Low	Low	Low	High	High
Overhead	High	High	Low	Low	Low	High

2) *Network mobility management*

NEMO (network mobility) was introduced in 2005 for network mobility problems [11]. As base station is not directly accessed by all users, as mobile host can only be accessed by using mobile routers (MR). Mobile router has its own home address. When the MR moves to a foreign access router it requires Care Of Address (COA) from the visited network. When it get its COA it sends the update message to it's HA (Home Address). Then HA of the MR forward this message to all data packets. The network mobility solutions like NEMO leads to reduced handoff, scalability, reduced complexity.

III. RELATED WORK

Considering the importance of mobility management in VANET researchers have suggested various techniques for mobility management as mentioned below:

Hayong Oh, Joon Yoo, Chang-kwon Kim and Sang hyun Ahn [13] in 2009 suggested a handover scheme to support multimedia services in VANET as well as in Vehicular Intelligent Transportation System. It uses FMIPv6 which reduces MIPv6 handover latency by using handover prediction, but it fails to manage sudden direction change of vehicles. It reduces handover latency using DAD (Duplicate Address Detection) process which helps to limit IP configuration delay. This model compares the handover delay of proposed scheme with FMIPv6 by using parameters like mobile position in AR and handover latency. This leads to robust handover by preventing original COA.

Ravi Shankar Shukla, Neeraj Tyagi [11] in September 2013 suggested a network mobility approach in VANET. In this model the movement of vehicles from one network to other is described. Every vehicle is equipped with MR, which is connected with AR. This model proposes that handover taking place at MR1 can also use MR2 for internet connectivity until handover is completed. This model supports seamless mobility of vehicles connected to mobile network across heterogeneous network in VANET by no service disruption during handover process between different ISPs using AR and MR.

Muhammad Nawaz Khan, Ishtiaq Wahid, Gulzar Mehmood [5] in 2012 proposed a handover scheme based on HMIPv6 and FMIPv6 for fast mobility of the vehicles and small coverage range of 802.11 Access Point (APs). In 802.11 Access Point due to fast mobility frequent handoff takes place, as a result throughput reduces. Mobility in VANET supports MIPv6 which leads to issues like packet loss, triangular routing and greater latency. So handover scheme based on FMIPv6 and HMIPv6 reduces handover latency by eliminating the DAD (Duplicate Address Detection) process. Also it minimizes the delay resulting due to binding update messages by ensuring unique IP address allocation.

Jong-Tae Park, Seung-Man Chun, Jun-Hyuk Choi and Seung-Mu Lee [12] in 2012 suggested SMMP. Simple Mobility Management Protocol (SMMP) provides global seamless handover in both homogenous and heterogeneous network. MMIPv6 is not used here but Session Initiation Protocol (SIP) is used here. It uses the concept of location management for mobility management by transmission of packets by bi-directional tunnel between mobile hosts. SMMP is compared with HMIPv6 and MIPv6 by using parameters like Pear signal ratio, Packet loss, Handover delay which concludes SMMP provides globally seamless IP handover.

S. Cespedes, X. Serman Shen [7] in 2010 suggested a scheme for seamless internet access in urban Vehicular Scenarios by using Hybrid HIP-PMIPv6. The proposed scheme is compared with NEMO, MIP and HIP. Analysis is done on the mobility handover delay and packets dropped which results to give better results on using Hybrid HIP-PMIPv6.

Dae Won Lee, Yoon-Ho Kim and Hwa Min Lee [14] in 2014 proposed a scheme which focuses on network mobility management to provide reliable communication within the vehicle even in highway scenario. This proposed mobility management scheme provides robustness of frequent path disruption caused by frequent vehicle mobility this leads to reducing handoff latency, minimizing packet loss during transition and reducing the signaling overhead by base unit and provides transparency to corresponding host.

Rodolfo I. Menguette, Luiz F. Bitten Court and Edmundo R.M. Mdeir [16] in 2012 proposed A Seamless Flow Mobility Management Architecture to provide good quality of service for vehicular applications with network based mobility management. Using NS3 simulator on different scenarios is tested to analyze the behavior of this architecture. The proposed architecture deals with different network interfaces at the same time to provide maximum network throughput, to decrease the handover time and to satisfy maximum number of packet loss and latency for each class of vehicular network application.

Jerome Harri, Christian Bonnet [4] in 2007 suggested a location aware framework called Kinetic Graph by using ad hoc protocols to implement reactive mobility management. Frameworks are suggested to kinetic graph which are beneficial for location awareness.

D. Rajini Girinath, S. Selvin [15] in 2010 suggested a cluster based routing algorithm named as Location based Multipath Flooding for hybrid mobility model to regulate the vehicular traffic. This algorithm helps to transmit real time updated information and maintain long link duration using NS2 simulator. The combination of clustering and routing patterns does influence the performance of VANET mobility.

Various authors have covered various mobility management techniques and their issues in the survey as listed above. Table given below summarizes these issues.

TABLE II MOBILITY MANAGEMENT ISSUES

Paper Name	Issues Covered	Remarks
Mobility and Handoff Management in VANET[2009]	Various host and network mobility solutions	Solution to mobility issues is to use NEMO to provide IP mobility support and various host specific protocols to be employed in more realistic scenarios.
An efficient Hybrid HIP-PMIPv6 scheme for seamless internet access in urban vehicular scenario[2010]	Interworking HIP-PMIPv6 scheme.	Mobility issues need improvement. Author suggests the need of Proxy Mobile IPv6 (PMIP) and further extensions.
Seamless quality driven multi hop data delivery scheme for video streaming in urban VANET scenario.[2011]	Multi hop PMIPv6 for VANET	For multi hop authentication problem in PMIPv6 author propose EM3A, a authentication scheme that guarantees authenticity of both MN and RN.
IP mobility management for VANET: challenges and solutions [2011]	Evaluation of NEMO RO solutions of VANET.	Due to handover delays and packet drops various mobility issues arises. Use of PMIPv6, NEMO protocols can be beneficial.
Global mobility and handover management for heterogeneous network in VANET[2013]	Network mobility among vehicles	For highways issues author suggest more and more access routers (AR) through which MR are connected, which helps MR1 to use MR2 services when handover takes place
Different mobility and handover management techniques in VANET[2013]	Comparison among SIGMA, MMIPv6, SMMP, EAR-HMIPv6 with MMIPv6.	Author suggests that recent mobility schemes to be evaluated in more realistic scenarios as mostly use NS2-network simulators.

IV. Open Issues and Future Scope

Lot of work has been reported in the related work section number III related to mobility management and handover techniques in VANET. Based on this related work various open issues have been extracted.

In Ad hoc routing various limitations which leads to degrading handoff performance with increasing number of hops, for which authors have given the solutions but they are unable to resolve. Traditional mobility model are not suitable for performance evaluation as they assume a random direction and speed selection. Major advancements still have to be taken in the field of providing security. Mobility management has also problem of data access and address configuration. All requests to compete for same limited bandwidth and time constraint are other challenges. Also miss of data upload leads to data staleness. Merging of dedicated VANET technologies like RAN (e.g, WiMAX) and 802.11p are still to be done, which is a challenge. Proper balancing of high priority safety info and users demands for infotainment applications are still the challenges to be resolved.

Future scope comprises an improved evaluation of various recent techniques like MMIPv6, SMMP, HMIPv6 and many more in more realistic vehicular scenarios. Still now accurate mobility patterns, network models and network performance which are the parameters on which evaluation relies are not available yet. There is a need of improving old technologies by keeping in mind large number of vehicles, frequently changing topologies, highway scenarios and high speed of vehicles. Mobility management issues are increasing rapidly so to resolve these issues there is a lot of scope in this area. Researchers have done a lot of work. Some of the issues have been solved built still there are many unsolved issues.

V. Conclusion

VANET is a critical part of ITS, so it has attracted significant research interest as the requirement of mobility management techniques are increasing. In this paper a comprehensive survey of mobility management techniques for VANET have been presented. The mobility management solutions for vehicular networks based on V2I or V2V communication have also been described. Some of the existing work for V2V and V2I mobility management has also been reviewed. Some light have also been thrown on several open research issues. Types of mobility management like handoff management and location management have also been outlined. In this paper various issues like security, performance evaluation, data access, address configuration, broadcast and routing issues, collision warning have been also presented. It is hoped that the survey done in this paper will prove to be beneficial to researchers working in the area of mobility management and its issues in VANET.

Acknowledgment

I am very thankful to Almighty God; whose grace and blessed mercy enabled me to complete this work with full devotion.

I am also grateful to Dr. Jyoteesh Malhotra, Head of the Department, Department of Computer Science Engineering, GNDU RC Jalandhar, for their invaluable support and guidance throughout this research work. Then, I would like to thank to my family and friends for their encouragement. I am all thankful to all those who helped me in one way or other at every stage of my work.

References

[1] Akyildiz IF, McNair J, Ho JSM, Uzunalioglu H, Wang W "Mobility management in next-generation systems" in 2012.

[2] Bechlar M, Wolf L. Mobility management for vehicular ad hoc networks. In proceedings of VTC 2005 Spring, vol. 4, 2005.

[3] W. Alasmary and W. Zhuang, "Mobility impact in IEEE 802.11p infrastructureless vehicular ad hoc networks," Ad Hoc Networks, to appear.

[4] Kinetic Mobility Management Applied to Vehicular Ad Hoc Network Protocols By Jerome Harri, Christian Bonnet, Fethi Institute Eurecom Department Communications Mobiles in December 21 2007.

[5] A Global Mobility Management Scheme for Reducing Overall Handover Latency in IP based VANETs by Muhammad Nawaz Khan, Ishtiaq Wahid, Gulzar Mehmood Internal Journal of Ad Hoc, Sensor Computing (IJASUC) in February 2012.

[6] M. Asefi, S. Cespedes, X. (Sherman) Shen, J. Mark. "A Seamless Quality Driven Multi-hop Data Delivery scheme for Video Streaming in Urban VANET scenarios" in IEEE ICC 2011. Kyoto, Japan.

[7] S. Cespedes, X. (Sherman) Shen. "An Efficient Hybrid HIP-PMIPv6 Scheme for Seamless Internet Access in Urban Vehicular Scenarios", in Proc. GLOBECOM. 2010.

[8] Kun Zhu, Dusit Niyato, Ping Wang, Ekram Hossain and Dong In Kim " Mobility and Handoff Management in Vehicular Networks: A survey" in Wiley Inter- Science in 2009.

[9] Blum J, Eskandarian A, Hoffman L. "Mobilty management in IVC networks". In Proceedings of IEEE Intelligent Vehicles Symposium, 2003.

[10] Sallet H. Langar R, Basir O, Boutaba R. "Proposal and analysis of region based location service management protocols for VANET" in 2008.

[11] Ravi Shankar Shukla, Neeraj Tyagi. " Global Mobility And Handover Management for Heterogeneous Network in VANET" in September 2013.

[12] Jong-Tae Park, Seung- Man Chun, Jun-Hyuk choi and Seung-Mu Lee, "Simple Mobility Management Protocol for Global Seamless Handover", in 2012 IEEE.

[13] Hayong Oh, Joon Yoo, Chang-kwon Kim and Sang hyun Ahn, " A Novel Mobility Management for Seamless Handover in V2V/V2I Networks", in IEEE in 2009.

[14] Dae Won Lee, Yoon- Ho Kim and Hwa Min Lee, "Route prediction based Vehicular Mobility Management Scheme for VANET", in May 2014.

[15] D. Rajini Girinath and S. Selvan, "A novel Cluster based Routing Algorithm for Hybrid Mobility Model in VANET", in 2010.

[16] Rodolfo I. Menguette, Luiz F. Bitten Court and Edmundo R.M.Mderia, "A Seamless Flow Mobility Management Architecture for VANET" in 2012.

Survey on Various Security Attacks and the Mitigation Techniques for MANET

Ashima Mittal[1] and Satwinder Singh[2]
[1,2]Department of Computer Science,
Guru Nanak Dev University, Regional Campus, Jalandhar, India
E-mail: [1]er.ashimamittal2013@gmail.com, [2]er.satwindermehmi@gmail.com

***Abstract*—Mobile ad-hoc network is a network of Mobile nodes that are connected with a wireless Link. There is no centralized node that controls the Network. The recent trend in mobile ad-hoc networks is on-demand routing where routes are established on Demand. Various security threats come into existence these days. Mobile ad-hoc networks are highly dynamic in nature. So, secure routing is a major issue now days. This paper is a survey on various security attacks, various mitigation techniques proposed by various researchers protocols for secure routing and the research on current trends.**

Keywords: MANETs, Routing, Denial of service (DoS)

I. Introduction

Wireless communication is growing day by day due to its Increasing applications. In recent years, MANETs (mobile ad-hoc Networks) have received more attention due to its self-creation and self-maintenance nature. Each device in MANET is free to move in any direction which results the change in link table frequently. The member nodes are itself responsible for all the link management. Each node in a MANET has its own wireless transmitter and receiver so that nodes can communicate with each other in their wireless range. The nodes which are not within the wireless range communicate with other nodes hop by hop by following some rules known as routing protocols. The latest work is done in wireless technology achieve a lot of attention. An ad-hoc network is one of such advancement in wireless technology which gives a new platform to wireless self-organized networks. The ad-hoc networks are not infrastructure networks and create routes when required. They are peer-to-peer network. They are mainly used for military oriented purposes. Confidentiality, integrity, availability, non-repudiation and authentication are the basic requirements of information security [2]. The dynamic nature of mobile ad-hoc networks creates a problem in finding multi-hop routes for communication path. In ad-hoc networks mobile node can move randomly because each node act as a router, so it is very difficult to find an optimal route. Security is still the main topic for many researchers. They provide various security routing protocols for secure communication.

II. Various Routing Protocols

In MANETs, some form of routing protocol is required in order to dynamically detect the multi-hop paths through which packets can be sent from one node to another [1]. Basically, there are two types of routing protocols for MANETs.

A. *Table Driven (Proactive)*

In this, every node maintains a table that represents the entire network topology. Various proactive protocols are DSDV, GSR and WRP.

B. *On-Demand (Reactive)*

In this, routes are not predefined for routing. Here, The Source node initiates the route discovery procedure when needed. Some reactive protocols are ABR, AODV, DSR and LAR. Much of the research has been done working in efficiency and stable routing of the MANETs. Because of the previous research, now we have huge number of routing protocols that are marvelous in terms of efficiency. But the situation changes when we focus on the security requirements of these protocols. The more detailed research is under the way to find a secure routing protocols. Due to the Open medium communication and dynamically changing behavior of MANETs, they are more vulnerable to attacks. Researchers have provided various secure routing protocols by watching their resistance towards various security attacks. Various proposes protocols by researchers are Secure Efficient Distance Vector Routing SEAD), Authenticated Routing for Ad-hoc Networks (ARAN), Secure Routing protocol (SRP) and many more.

III. Some Security Attacks

There exist some potential loopholes in MANETs that can be exploited by undesirable nodes to destroy the smooth functioning in the network. In MANETs, attacks are classified into two types: active attacks and passive attack. Brief introduction about both the attacks is given below.

A. *Passive Attacks*

In this attack intruder snoop packet that contain secret information by listening only to the channel e.g. IP addresses, location of nodes etc., without disturbing the operation of network. These attacks are very difficult to identify.

1. *Eavesdropping:* In this attack, the malicious node obtain some secret information e.g. password of the node that is very important to kept secret, location, private key and public key.

2. *Traffic Analysis:* In this attack, attacker detects transmission to inflict important information such as source-destination pair.
3. *Jellyfish Attack:* In this attack, attacker breakdown the performance of the network by introduces the delay in sending packets that it receives.

B. Active Attacks

In active attacks, malicious nodes confuse the network topology by introducing false information to it. They can do two things either attract the traffic or compromise the packets. They can send packet to the wrong node.

1. *Denial of Service (DoS) Attack:* In this, attacker does not corrupt data; he can just disable services by replacing them with the virtual services.
2. *Wormhole Attack:* An attacker took the packet from one location in the network, tunnels the packet to another location and again resends the packet in the network but at different location. Wormhole attack is a big threat to security of MANETs. The wormhole attack can be detected and prevented by implementing digital signature.
3. *Sinkhole Attacks:* In this, a sinkhole node becomes the attraction point for all the nodes and attracts the data toward itself from other neighboring nodes. Sinkhole node maintains the route according to itself, create complicated network and finally destroy the network.
4. *Gray-Hole-Attack:* In this, malicious node behaves like it is an actual node during the discovery process. After route discovery process when sender sends the packet it silently drop the packets sent to it.
5. *Fabrication Attack:* It is an active attack which breaks the network authenticity by acting like it is a source node. It then sends an error message to the nodes to inform that the network is no more exists. Other nodes update their table with false information. In this way it drops the routing performance of a network.

IV. Related Work on Various Security Techniques

The Authors in [3] presented a design and performance evaluation of new on-demand ad hoc network routing protocol known as Ariadne. Ariadne helps the protocol by preventing attacker from altering with uncompromised routes consisting of such uncompromised nodes. Ariadne also helps to prevent Denial-of-Service attacks. Some more features of Ariadne is that it is efficient and using only efficient symmetric cryptographic operations. They also compared Ariadne to a version of Dynamic source routing (DSR) by disabling all protocol optimizations that are not present in Ariadne and then calculate the effect of optimization and security separately. They prove that Ariadne lowers the packet overhead by 41% than for unoptimized DSR. However Ariadne added some cost for security that was not present on unoptimized DSR.

Cheng Yong, Huang Chuanhe and Shi Wenming in 2007 suggested novel secure routing protocol for mobile ad-hoc networks known as trusted dynamic source routing (TDSR) [4]. In this a trust score is calculated on the basis of direct trust and indirect trust. When the trust value of the node falls below the threshold then it is added to the blacklist. The nodes that performs below the threshold or present in blacklist are not b forwarded.

Dhurandher and Mehra in 2009 [5] introduced the approach that can be used to calculate the trust value of node in a dynamic manner and also protects message modification by attacker. The result is calculated by doing simulations in packet delivery ratio and the number of times packet was broken into parts. By considering behavior of a node a trust value is given to a node. it can be incremented and decremented according to the behavior of node. Trust value can be of three types that are: positive, negative or zero that shows that node is known, malicious or unknown behavior respectively.

Pallavi and Trivedi in 2011[6] gave solution to prevent serious attack that is a wormhole attack by the use of digital signatures. In this if a sender wants to send packet to destination node it will create a secure path with the help of digital signature verification. Node sends a packet along with a digital signature and if it matched with the digital signature stored in their database of other nodes then the request is from authentic source.

Kamini Nalavade and Dr. B.B. Meshram in June, 2014 gave the layered approach for preprocessing of data in intrusion detection system [7]. To remove unwanted and redundant data from packets, the layered approach of TCP/IP model is used for the faster preprocessing of data in intrusion detection system.

S. Saravanakumar, Umamaheshwari, D. Jayalakshmi and R. Sugumar [8] in 2010 handles the issue of complexity and throughput that are the problems in Intrusion Detection System (IDS). The authors compare various IDS systems and then suggests a scheme that uses the combination of artificial neural network algorithms. This combination of algorithm gives better performance.

Jiwen CAI, Ping YI, Jialin CHEN, Zhiyang WANG, Ning LIU [9] in 2010 proposed the algorithm to detect black hole and gray hole attacks in adhoc networks. The researchers demonstrates the adaptive approach using cross layer design. The authors proved their theory by using path-based method to overhear the next node. So, it saves system resources by not sending out extra control messages. A collision rate reporting system is established to reduce the false positive rate under high network load.

TABLE 1 A BRIEF SUMMARY OF VARIOUS KINDS OF SECURITY ATTACKS ON MANETS AND THEIR MITIGATION TECHNIQUES

Security Attack	A Brief Description of the Security Attack	Techniques Proposed to Mitigate the Attacks
1. Black-Hole Attack	Also called Packet drop attack in which intruder drop all the packets advances to it.	1. The technique in [12] makes use of creditable routing table to detect the black-holes and eliminate them. 2. The authors in [9] verify the control messages sent by the attacker and check if the black-hole attack is executed. 3. In [13] a two-stage technique is followed. In the first stage detection of the malicious nodes is done and in the second stage the removal of malicious node is done.
2. Wormhole Attack	In this attack, an intruder records the packets at one location and forwards them to another location.	1. The wormhole attack can be detected and prevented by implementing digital signature. This is proposed by [6]. 2. In [14] the researchers use a cooperative approach among the distributed nodes to detect and minify the wormhole attack.
3. Sink-Hole Attack	In a Sink-Hole attack, the intruder node/malicious node sends fake routing information claiming that it has an optimum route to the target which causes other nodes in the Ad Hoc Network to route data packets through it.	1. A trust based algorithm is implemented in [15] to diminish sink-hole attack. 2. In [10] they make use of technique in which the mobile agents are used to detect and reduce the sink-hole technique. 3. The authors in [16] implement scheme which include three variables: Sequence Number, Route Add Ratio and Previous Image Ratio to prevent and mitigate the Sink-hole Attacks.
4. Grey hole attack	It drops the part of the data and cheats the previous node.	1. In [17], author proposed mechanism to detect gray hole attack. The detection involves proactively invoking of collaborative and distributive algorithm involving neighbors. The detection decision is based on threshold cryptography.
5. Route Fabrication	In this attack the wrong routing massages are sent into the network by the intruder.	1. In [18] the authors use fuzzy logic, a soft computing method to establish a quantifiable trust value among the nodes of the network. This approach prevents the route fabrication attack.

QuanJia, Kun Sun and Angelos Stavrou [11] in 2011 designed an approach to prevent Denial of Service (DoS) attack. This approach is designed for multipath communication in mobile ad hoc networks (MANETs). They defined the capability messages that are exchanged in between the nodes of network. This enables the each node to maintain overall throughput of flows in the network and then dynamically adjust local constraints. it helps to prevents DoS attacks against a specific node.

Following Table I show the brief survey on different kind of security attacks and some mitigation techniques to minify the effects of these attacks.

V. RESEARCH GAPS

As we all know MANETs are always very attractive for the military purposes and lot of research is going on this topic. Security is always a major topic in research field. As MANETs use is increasing day by day, new attacks are also coming forth continuously. So, the main research gap in the field of security is that researchers are only focusing on some nodes and some routing protocols. None of the existing system is a complete solution for the security attacks. As we have limited resources in MANETs, this topic is also less explored. Researchers should examine more security risks to explore the topic of security and also to find its solution.

VI. CONCLUSION

This paper gave all the stock information about the security of ad hoc networks. In the introduction section we discussed about the MANETs. We also discussed about the routing protocols and its types. In the next part, we discussed some of the main security attacks that are vulnerable to ad hoc networks. This paper proposed the related work on the security threat by many researchers and the research gap in this field. Lot of work is going on the security attacks by intruder.This paper is a survey on various methods that are proposed by researchers to prevent security attacks and the what the researchers should more focus about security of MANETs.

ACKNOWLEDGMENT

We would like to thank almighty for his constant blessings. We would also like to thank Dr. Jyoteesh Malhotrawho provide valuable advices and our friends for helping and supporting us throughout the making of this paper.

REFERENCES

[1] Jayraj Singh, Arunesh Singh, Raj Shree “An Assessment of Frequently Adopted Security Patterns in Mobile Ad hoc.

[2] Stallings W [2000], Network Security Essentials: Security Attacks. Prentice Hall. (pp. 2-17).

[3] YIH-CHUN HU∗and ADRIAN PERRIG Carnegie Mellon University, USA DAVID B. JOHNSONRice University, USA.in 2005.

[4] CHENG Yong, HUANG Chuanhe, SHI Wenming, “Trusted Dynamic Source Routing Protocol”, Wireless Communications, International Conference on Networking and Mobile Computing, WiCom2007, Sept. 21-25,2007,pp.1632-1636.

[5] Sanjay K. Dhurandher, VijetaMehra, “Multi-path and Message Trust-Based Secure Routing in Ad Hoc Networks”, International Conference on Advances in Computing, Control, andTelecommunication Technologies, ACT '09. Dec. 28-29,2009,pp.189-194.

[6] Pallavi Sharma, Aditya Trivedi, “An Approach to Defend Against Wormhole Attack in Ad Hoc Network Using Digital Signature”, 2011 IEEE 3rd International Conference on Communication Software and Networks (ICCSN), May 27-29 2011, pp.307-311.

[7] International Journal of Computer Applications Technology and Research (IJCATR) Volume 3 Issue 6 June 2014 Layered Approach for Preprocessing of Data in Intrusion Prevention SystemsKamini Nalavade,Dr. B. B. Meshram.

[8] S. Saravanakumar, Umamaheshwari, D. Jayalakshmi, R. Sugumar, “Development and implementation of artificial neural networks for intrusion detection in computer network”, Int. Journal of Computer Science and Network Security2010. vol. 10, no. 7, pp. 271-275.

[9] Jiwen CAI, Ping YI, Jialin CHEN, Zhiyang WANG, Ning LIU, “An Adaptive Approach to Detecting Black and GrayHole Attacks in Ad Hoc Network”, 2010 24th IEEE International Conference on Advanced Information Networking and Applications (AINA),Perth, Australia,April 20-23, 2010, pp.775-780.

[10] D.Sheela, Naveen Kumar. C, G.Mahadevan, “A Non-Cryptographic method of Sink HoleAttack Detection in WirelessSensor Networks”, 2011 International Conference on Recent Trends in InformationTechnology(ICRTIT),Chennai, India,June 3-5, 2011, pp.527-532.

[11] Quan Jia, Kun Sun, Angelos Stavrou, “CapMan:Capability-based Defense against Multi-Path Denial of Service (DoS) Attacks in MANET”, Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN), Maui, HI, USA 2011, July 31-August 4, 2011, pp.1-6.

[12] Japing Wang, Haoshan Shi, “A Secure DSR Protocol Based on the Request Sequence-Number”, 5th International Conference on Wireless Communications, Networking and Mobile Computing, 2009 (WiCom '09), Beijing, China, Sept. 24-26, 2009, pp. 1-4.

[13] Muhammad Zeshan, Shoab A. Khan, Ahmad Raza Cheema, Attique Ahmed, “Adding Security against Packet Dropping Attack in Mobile Ad hoc Networks”, International Seminar on Future Information Technology and Management Engineering, (FITME '08), Leicestershire, UK, Nov. 20. 2008, pp.568-572.

[14] Gunhee Lee, Dong-kyoo Kim, Jungtaek Seo, “An Approach to Mitigate Wormhole Attack in Wireless Ad Hoc Networks”, International Conference on Information Security and Assurance (ISA 2008), April 24-26, 2008, pp.220-225.

[15] Thanachai Thumthawatworn, Tapanan Yeophantong, Punthep Sirikriengkrai, “Adaptive Sinkhole Detection on Wireless Ad Hoc Networks”, Proceedings of IEEE Aerospace Conference, 2006,Big Sky, Montana, USA, 4-11 March 2006,pp.1-10.

[16] Benjamin J. Culpepper, H. Chris Tseng, “Sinkhole Intrusion Indicators in DSR MANETs”, Proceedings of First International Conference on Broadband Networks(BroadNets 2004),San Jose, USA, Oct. 25-29, 2004, pp. 681- 688.

[17] Jaydip sen et. al “A Mechanism for Detection of Gray Hole Attack in Mobile AD Hoc Networks” ICICS 2007, IEEE.

[18] H. Hallani, S.A. Shahrestani, “Trust Assessment in Wireless Ad-hoc Networks”, Wireless Days, 2008 (WD '08). 1st IFIP, Dubai, Nov. 24-27, 2008, pp.1-5.

Advance Stable Election Protocol in Wireless Sensor Networks

Mandeep Singh[1] and Navjot Sidhu[2]
[1,2]*Centre for Computer Science and Technology,*
Central University of Punjab, Bathinda, India
E-mail: [1]bhadour.cup@gmail.com, [2]navjotsidhu8@gmail.com

***Abstract*—Wireless sensor network (WSN) is an emerging research field. There are large numbers of sensors that collect and send data to base station. Saving energy by using various routing techniques is a challenge. Clustering is main technique used for this. Various protocols like SEP (Stable Election Protocol) and ESEP (Extended Stable Election Protocol) are clustering based heterogeneous aware protocols. In this paper, a new protocol ASEP (Advance Stable Election Protocol) has been proposed based on SEP. This is based on changing more efficiently and dynamically the cluster head election probability. Performance of this protocol has been evaluated in MATLAB and graphical results have been shown. The performance of ASEP is better than SEP in form of first node dies and total number of packets delivered.**

Keywords: WSN, Clustering, LEACH, SEP, ASEP

I. Introduction

Wireless Sensor Networks (WSNs) have been considered as the very important technologies used in this world from last two to three decades. There is lot of advances in wireless communication technologies that gives us the small, smart and cost effective sensors. These sensors can be deployed anywhere and connected through wireless links [1].

There is large number of sensor nodes in wireless sensor network. These can be deployed inside any field or near that field where the sensing has to be done. A sensor node mainly consists of five components:

1. Sensing
2. Memory
3. Processor
4. Transceiver (transmitter and receiver)
5. Battery [5]

In wireless sensor network (WSN), there are number of sensor nodes that send their data to the sink or base station. Base station is a powerful node or device that can receive the data from the sensor nodes. If the base station or sink is in the radio range of the node then data can be send directly otherwise the other sensor nodes can be used [3].

II. Routing in Wireless Sensor Networks

Energy is the first and foremost consideration for the routing protocols used in Wireless Sensor Networks. According to the application of the network routing protocols can be different [8].

The characteristics of wireless sensor network are different from other networks like cellular networks or Mobile Ad-Hoc networks. So, routing process is also different. As we know that sensor network have large number of sensor nodes, so traditional IP based schemes cannot be applied here [7],[8].

III. Hierarchical Networks Routing

The major design attribute of wireless sensor networks is Scalability, similar to other communication networks. As we know that there is large number of sensor nodes. So, single tier can cause latency due to overload of Gateway. Due to this and wide area covered by nodes, long distance communication is difficult and single tier network does not work efficiently. Here comes the concept of clustering. In clustering, network is divided into clusters. [8]. There is Concept of cluster head used to reduce the work done by normal sensor nodes. Normal nodes sense the data to cluster heads and cluster heads aggregate the data and send it to base station. [7]. There are various protocols that use the concept of clustering. Like LEACH, SEP, ESEP, HEED, DEEC, etc. Our work is based on SEP (Stable Election Protocol) which is a heterogeneous aware protocol. Heterogeneous means nodes have not same initial energy. Some nodes have more initial than other nodes in these protocols. The detail of SEP is given below.

A. Stable Election Protocol

Stable Election Protocol uses the basic techniques of the LEACH protocol like cluster hierarchy, choosing optimal number of clusters, energy model used and optimal probability of a node to become the cluster head. In SEP, nodes are heterogeneous in nature means nodes have not same initial energy. SEP have the fraction of advanced nodes (m) (nodes which have more energy than the normal nodes, where m is the percentage of advance nodes in total nodes) and the additional energy factor between advanced and normal nodes (α). In this, advanced nodes have to become cluster heads more often than the normal nodes. This new heterogeneous setting (with normal and advanced nodes) has no effect on the spatial density of the network. But, the total energy of the system changes. Suppose if E_0 is the energy of each normal sensor node. The energy of each advanced sensor node will be $E_0 \cdot$

(1 + α). The total energy of this new heterogeneous setting is equal to: [11]

$$n\,(1 - m)\,E_0 + n\,m\,E_0 \cdot (1 + \alpha) = n\,E_0\,(1 + \alpha\,m) \quad (1)$$

where, n= No. of nodes in sensor network

α= Energy factor between advance and normal nodes

m=Percentage of advance nodes in network

So, the total energy of the system is increased by (1+*α.m)* times

According to the additional energy in the advance nodes, the probability density function of the protocol also changes. The probability of various nodes to become cluster heads is given

$$p_{nrm} = \frac{p_{opt}}{1+\alpha.m} \quad (2)$$

$$p_{adv} = \frac{p_{opt}}{1+\alpha.m} \times (1+\alpha) \quad (3)$$

Where p_{nrm} is the probability of the normal nodes to becomes the cluster head and p_{adv} is the probability of the advance nodes to becomes the cluster head [11].

As now it is known that advance nodes with more initial energy have more chances to be a cluster head. So it increases the network lifetime due to its heterogeneous aware algorithm. But after some time, when the energy of the advanced and normal nodes remains same, it again chooses advance nodes increasing times than normal nodes. Due to this, energy of advanced nodes depletes more quickly than normal nodes. It decreases lifetime of network. So ASEP (advance stable election protocol) has been proposed in which advance nodes are selected as cluster head in starting rounds, when energy depletes upto a certain level after that the probability of both advance and normal nodes will be same.

B. *Advance Stable Election Protocol the Details of ASEP is given Below*

1) *Heterogeneous WSN model used*

In this section, the model of wireless sensor networks will be described. Nodes are heterogeneous in the initial amount of energy. Particularly, the setting of Wireless Sensor Networks and energy model used is presented. It also tells us that how the optimal clusters can be computed. There is some percentage of nodes with extra amount of energy than other nodes. There are p and k percentage of nodes equipped with a and b times more energy than other nodes. These nodes are called advance and moderate nodes. All these sensor nodes are distributed uniformly over the sensor field [11][12].

2) *Creation of a cluster*

In this setting, wireless sensor network is hierarchically clustered. LEACH (Low Energy Adaptive Clustering Hierarchy) is a protocol that uses same type of clustering hierarchy. In LEACH, the concept is repeated in each round. Round has two phases: cluster creation phase and steady state phase. clusters are re-established in each round and new cluster heads are chosen in each round. Due to this, load is well distributed among the nodes of the network. In this type of cluster hierarchy, each node transmits to closest cluster head and this cost of communication is very-very less than the cost of sending directly to the base station. Only the head nodes communicate directly with the base station. This expends large amount of energy but each node do this periodically. In LEACH protocol, there is optimal percentage of nodes (p_{opt}) from total nodes that has to become cluster head in a round. This percentage is determined priori and it is assumed that nodes are distributed uniformly in the field.

In LEACH, nodes are homogeneous. It means that all the nodes have same initial energy. It guarantees that each and every node will become cluster head exactly once in every $\frac{1}{\mathrm{p_{opt}}}$ rounds. It is also assumed as the epoch of the network.

In start, every node has the probability equal to $\frac{1}{\mathrm{p_{opt}}}$ to become cluster head. In a epoch, when one node becomes cluster head once, it can-not become cluster head in same epoch. All the nodes that has not became cluster head yet belongs to set G. After each round in same epoch, the probability of each node increases which belongs to set G. each node belongs to set G chooses a random number between 0 and 1. If less than a threshold value, this node becomes cluster head in that round. This threshold equation is

$$\mathrm{T(n)} = \frac{\mathrm{p_{opt}}}{1-\mathrm{p_{opt}}\left(\mathrm{r\ mod}\left(\frac{1}{\mathrm{p_{opt}}}\right)\right)} \quad \text{if n } \in \mathrm{G} \quad (4)$$

Here r is the current number .at the end of the round, all nodes send data to cluster heads [11][12].

C. *Optimal Number of Clusters*

In all the previous protocols, nodes are assumed uniformly distributed and the optimal probability of the node to become cluster head is the function of spatial density. The optimal clustering mainly depends on the energy model used. Same energy model is used as in LEACH and SEP protocol.

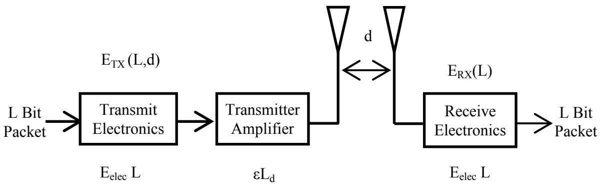


Fig. 1 Radio Energy Model Used [11]

As shown in above figure's radio energy dissipation model, in transmitting an L bit message over a distance d and to achieve an acceptable signal-to-noise ratio, the energy expended by radio is given by:

$$E_{Tx}(l,d) = \begin{cases} L \cdot E_{elec} + L \cdot \epsilon_{fs} \cdot d^2, if\ d < d_0 \\ L \cdot E_{elec} + L \cdot \epsilon_{mp} \cdot d^4, if\ d \geq d_0 \end{cases} \quad (5)$$

Where d is the distance between sender and receiver node, L is the size of the packet, E_{elec} is the energy dissipated per bit to run the receiver circuit or transmitter, ϵ_{fs} and ϵ_{mp} depend on the transmitter amplifier model used. When the equation given above is equated at d = d_0, the value comes $d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$. So to receive an L bit size message, it expends the energy as $E_{Rx} = L \cdot E_{elec}$. By taking all the basic assumptions and if the distance of any node to the base station or sink <= d_0 then the energy dissipated in a round by the cluster head node is given below:

$$E_{CH} = L \cdot E_{elec}\left(\frac{n}{k} - 1\right) + L \cdot E_{DA}\frac{n}{k} + L \cdot E_{elec} + L \cdot \epsilon_{fs} \cdot d_{toBS}^2 \quad (6)$$

Where d_{toBS} is the distance from sink or base station to base station, k is the total number of clusters and E_{DA} is the aggregation cost of a data bit per signal. The energy used by a non cluster head node is given below:

$$E_{nonCH} = L \cdot E_{elec} + L \cdot \epsilon_{fs} \cdot d_{toCH}^2 \quad (7)$$

Where d_{toCH} is the distance from any cluster node to its cluster head. So the total energy dissipated in a single cluster per round is given below:

$$E_{cluster} \approx E_{CH} + \frac{n}{k} E_{nonCH} \quad (8)$$

So the total energy dissipated in the network is given below:

$$E_{tot} = L \cdot \left(2nE_{elec} + nE_{DA} + \epsilon_{fs}\left(k.d_{toBS}^2 + n\frac{M^2}{2 \cdot \pi \cdot k}\right)\right) \quad (9)$$

When differentiation of E_{tot} has done with respect to k and result is equated to zero, the optimum clusters constructed comes that are given below:

$$k_{opt} = \sqrt{\frac{n}{2\pi}}\frac{M}{d_{toBS}} = \sqrt{\frac{n}{2\pi}}\frac{2}{0.765} \quad (10)$$

Where the average distance from base station or sink to any cluster head is given below :

$$E[\,d_{toBS}\,] = \int_A \sqrt{x^2 + y^2}\frac{1}{A}dA = 0.765\frac{M}{2} \quad (11)$$

So p_{opt}, the optimal probability of a node to become cluster head is given below:

$$p_{opt} = \frac{k_{opt}}{n} \quad (12)$$

In the LEACH protocol this is shown that if these clusters are not made by optimal way i.e. more or less than this optimal then the consumption of energy increases exponentially [11],[12].

The working of ASEP and SEP is same and only difference is in the probability density function. As it is clear from the equations (2) and (3) that advance nodes have more probability in the whole lifetime of the network. It surely increases the lifetime of the network than LEACH as the advance nodes are selected cluster head more often than normal nodes. There comes a time when there remains same energy in both advance and normal nodes but the advance nodes are selected as cluster head with same probability as before. It drains the energy of advance nodes more quickly than normal nodes. so the probability density function for advance nodes must be changed after attaining that particular energy level. To calculate that value, energy dissipated by both advance nodes and normal must be calculated, that is given below

3) How to calculate that energy value

The probability of advance nodes is higher, so there is possibility that advance nodes can become the cluster head in each round. So energy decreased from each advance node in a round is given below:

$$E_{AN} = L\left[\frac{n}{k} \cdot (E_{elec} + E_{DA}) + \epsilon_{fs} \cdot d_{toBS}^2\right] \quad (13)$$

Here $\mathrm{E_{CH}}$ is the energy that is dissipated by each advance node in a round. Then number of rounds possible for a cluster head with initial energy equal to $(1 + \alpha)\ \mathrm{E_0}$

$$\mathrm{NR_{AN}} = (1 + \alpha)\mathrm{E_0}/\mathrm{E_{AN}} \quad (14)$$

By this method, energy dissipated by normal node in a round can also calculated

$$E_{NN}{=}L\left[E_{elec} + \epsilon_{fs} \cdot d_{toCH}^2\right] \quad (15)$$

Number of rounds possible for a normal node is

$$\mathrm{NR_{NN}} = \mathrm{E_0}/\mathrm{E_{NN}} \quad (16)$$

After some rounds, advance and normal nodes will have same residual energy but the probability of advance nodes will have more than the normal nodes. so it will be selected more times than normal nodes and energy of advance nodes drain more quickly than normal nodes. there must be change in probability density function after nodes dissipate a particular amount of energy. this value of energy can be calculated by equation given below [13]

$$\mathrm{E_{level}} = \mathrm{E_0}\left(1 + \frac{\alpha\mathrm{E_{NN}}}{\mathrm{E_{NN}} - \mathrm{E_{AN}}}\right) \quad (17)$$

Probabilty Density Function of ASEP

$p_{nrm} = \frac{p_{opt}}{1+\alpha.m}$ for nml nodes, energy > E_{level} (18)

$p_{adv} = \frac{p_{opt}}{1+\alpha.m} \times (1 + \alpha)$ for advance nodes, energy > E_{level} (19)

$p_{an} = d\frac{p_{opt}}{1+\alpha.m} \times (1 + \alpha)$ for nml, adv nodes, energy < E_{level} (20)

Where E_{level} = G. E_0

It is assumed that all advance nodes will be cluster heads in all rounds but in reality it is not the case. Also normal nodes will become cluster heads for some rounds. So the exact value of G cannot be find, but through various simulations with random topologies, nearest value of G can be found. For the best result of protocol in terms of first node die and total packets delivered to base station, the value of G has come 0.7.

d is the variable which is used to control the cluster head number. If the number of cluster head will be more then the nodes will lose their energy very rapidly and die. If the number of cluster heads will be less then it will be difficult to gather data from the field. Through various simulations with random topologies, the nearest value of d can be found. For the best results of d in terms of first node die and total packets delivered, the value of d has come 0.05.

IV. SIMULATION AND RESULTS

A. *Simulation Parameters*

These are the basic parameters used for simulation. The main parameters are sensor area, base station position, number of nodes used in simulation, initial energy of the normal nodes, optimal probability of node to be cluster head and data aggregation cost [11].

TABLE I SIMULATION PARAMETERS

Parameters	Values
Sensor Field	(100 × 100)
Sink Position	(50, 50)
N	100
Packet Size	4000bits
ε_{fs}	10pj/bit/m^2
ε_{mp}	0.0013pj/bit/m^4
p_{opt}	0.10
E_{DA}	50nj/bit
E_0	0.5j
M	0.1
α	1
d_0	87.7m

A. *Performance Metrics*

1) *First node dies*

This parameter tells us about the round when first sensor node dies. It is also called the stable region of the network. If this region is long, then the nodes will send more useful information about the environment to be sensed to base station.

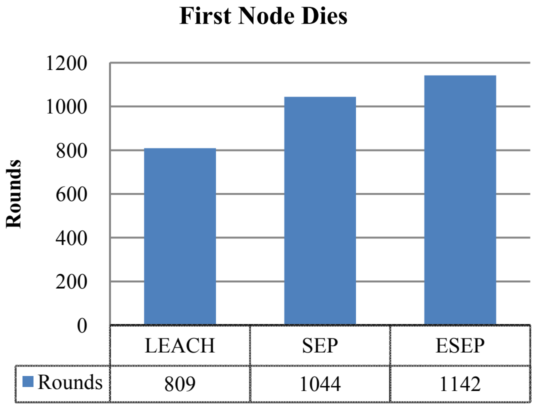


Fig. 2 First Node Dies

The stable region of the network in LEACH protocol is around 809 rounds. In SEP and ASEP, it is near 1044 and 1116. The performance of network is increasing from SEP to ASEP due to changes done in probability function of the SEP protocol.

2) *Total number of packets delivered*

The main work of the wireless sensor network is to send data to base station. This parameter tells us the number of packets that is sent to the base station from cluster heads. The performance of ASEP protocol is more than SEP protocol.

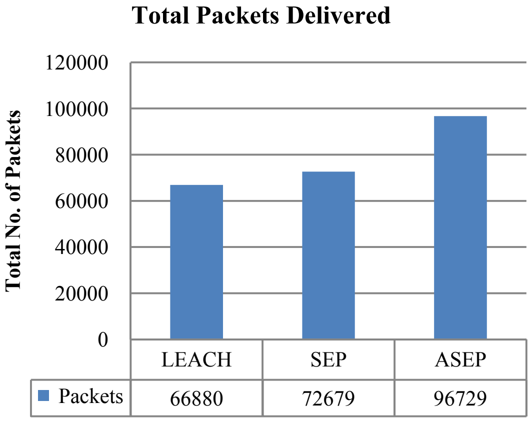


Fig. 3 Total Packets Delivered

CONCLUSION

Basics of wireless sensor network and heterogeneous aware protocol (Stable Election Protocol) have been discussed. The ASEP (Advance Stable Election Protocol) has been proposed which is based on SEP. it changes the probability function of the Stable Election Protocol and adds the point at which the energy of advance nodes remain same to the normal nodes. It increases the performance of network in terms of first node dies and total number of packets delivered to the base station.

REFERENCES

[1] J. Zheng and A. Jamalipour, " Introduction to Wireless Sensor Networks" in *Wireless Sensor Networks-A Networking Perspective*. Hoboken, New Jersey: John Wiley & Sons, pp.1-16., 2009.

[2] J. A. Stankovic, " Wireless Sensor Networks", *Computer,* vol. 41, no. 10, Pp. 92-95, 2008, DOI: 10.1109/MC.2008.441

[3] Chris Townsend, Steven Arms, "Wireless Sensor Networks: Principles and Applications" in *Sensor Technology Handbook*, Elsevier Inc., 2005 from http: //www.globalspec.com /reference /46556/203279/chapter-22-wireless-sensor-networks-principles and-applications

[4] I.F. Akyildiz, W. Sui, Y. Sankarasubramaniam, E. Cayirci, " Wireless sensor networks: a survey", *Computer Networks* 38, Elsevier Science B.V, pp. 393–422., 2002.

[5] L. Y. Min, W. S. Ci, N. X. Hong, " The Architecture and Characteristics of Wireless Sensor network", *Int. Conf. on Computer Technology and Development,* 2009, © IEEE DOI: 10.1109/ICCTD.2009.44., 2009.

[6] F. Martincic, L. Schwiebert, " Introduction to Wireless Sensor Networking", in *Handbook of Sensor Networks: Algorithms and Architectures*. Hoboken, New Jersey: John Wiley & Sons, pp.1-40., 2005.

[7] J. N. Al-Karaki, A. E. Kamal, " Routing Techniques in Wireless Sensor Networks: A Survey", *IEEE Wireless Communications archive.,* vol. 1, no. 6, Pp. 6-28., 2004.

[8] K. Akkayaand, M.Younis." A Survey on Routing Protocols for Wireless Sensor Networks", *Ad Hoc Networks, Elsevier B.V.* vol. 3, 2003, DOI:10.1016/j.adhoc.2003.09.010.

[9] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient routing protocols for wireless microsensor networks, " *in Proc. 33rd Hawaii Int. Conf. System Sciences (HICSS),* Maui, HI, Jan. 2000.

[10] W. R. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, " An application-specific protocol architecture for wireless microsensor networks", *IEEE Transactions on Wireless Communications*, 1(4):660–670, October 2002

[11] G. Smaragdakis, I. Matta, A. Bestavros, "SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks" *Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA 2004)*, 2004.

[12] M M Islam, M A Matin, T K Mondol, "Extended Stable Election Protocol (SEP) for Three-level Hierarchical Clustered Heterogeneous WSN", *IET Conference on Wireless Sensor Systems (WSS 2012)*, London, 18-19 June 2012.

[13] B. Elbhiri, R. Saadane, S. El Fkihi, D. Aboutajdine, "Developed Distributed Energy Efficient Clustering (DDEEC) for heterogeneous wireless sensor networks", *5th international Symposium I/V Communications and Mobile Network (ISVC)*, Rabat, 2010, DOI: 10.1109/ISVC.2010.5656252

Security Enhancement in AODV Protocol against Network Layer Attacks

Khushmeet Singh[1] and Amanpreet Kaur[2]
[1,2]Centre for Computer Science and Technology, Central University of Punjab, Bathinda, India
E-mail: [1]erkhushmeetsingh@gmail.com, [2]pandheraman@gmail.com

***Abstract*—A mobile Ad hoc network (MANET) is a wireless decentralized and self-configuring network in which nodes communicate with each other either directly or through intermediate nodes. Each node plays role of both transmitter and receiver. These unique features make MANET suitable for use in various emergency situations. Besides the various advantages, the open medium, rapidly changing topology and lack of centralized monitoring make MANET's vulnerable to different attacks. Hence it is vital to develop some security mechanism to protect MANET from attacks. In this paper performance of AODV Protocol is analyzed in presence of two network layer active attacks namely Blackhole and Malicious packet dropping. In active attack, an attacker disrupts the regular operation of the network, change the data or harm the system. Both these attacks are performed simultaneously on the network. Solution is proposed and simulated to prevent Blackhole attack. Malicious packet dropping attack is prevented by using watchdog intrusion detection system. Both solutions are combined and Performance of AODV and Modified AODV protocol is analysed under both attacks with respect to Packet Delivery Ratio, Average End-to-End Delay and Routing Overhead using NS2.**

Keywords: MANET, AODV, RREQ, RREP, CBR

I. INTRODUCTION

Today wireless networks are at its zenith. Every user wants wireless connectivity to communicate and transfer data with each other irrespective of their geographic position. Two main characteristics of wireless networks that propelled their widespread usage are mobility and ease of deployment. Laying cables in wired network is very time consuming and maintenance is also very high. Wireless communication today surrounds us in many colors and flavors, each with its unique frequency band, coverage, and range of applications. Among all the wireless networks, MANET is of its unique importance.

A mobile Ad hoc network (MANET) is a wireless decentralized self-configuring network in which nodes communicate with each other either directly or through intermediate nodes. Each node transmits and receives data through bidirectional links. All nodes are mobile and configure themselves in network without help of any infrastructure. Due to mobility, network topology changes over time because nodes join or leave the network at any time. These unique features (self-configuring, infrastructure-less, decentralized) make MANET suited for use in Emergency situations such as military operations, education, entertainment, sensor networks etc. Nodes may consist of broad range of devices like laptops, PCs, PDAs, smart phones as shown in Fig 1.

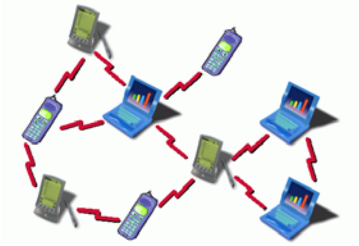


Fig. 1 Mobile Ad-hoc Network

Besides various advantages, the open medium, rapidly changing topology and lack of centralized monitoring make MANET vulnerable to various attacks that will be presented in later sections. In this case, it is vital to develop some security mechanism to protect MANET from attack. Security goals of MANET are same as compared to others networks i.e. Confidentiality, Integrity, Availability and Authenticity [10]. The rest of paper is organized as follows. Section 2 discusses various security goals in MANET. Section 3 discusses attacks in MANET. Section 4 describes the proposed solution to prevent Blackhole attack. Section 5 presents solution to prevent malicious packet dropping attack. Section 6 describes workflow of research. Section 7 presents the simulation set up and the metrics used to evaluate the performance. Results and discussion are shown in Section 8. Finally, paper is concluded in Section 9.

II. SECURITY GOALS IN MANET

Main aim of security services is to protect the data and the various resources from attacks. Various security goals in MANET are as follows [1]:

TABLE 1 SECURITY GOALS IN MANET

Security Goal	Purpose
Availability	Network services are always available whenever requested
Authenticity	Communication between nodes is legitimate
Data Confidentiality	Message exchange between two nodes cannot be understood by anyone else
Integrity	Message sent from sender node to receiver node was not modified by any malicious node during transmission
Non-Repudiation	Origin of the message is genuine. It guarantees that the sender and receiver of a message cannot deny later that it has not sent or receive the message

III. Attacks in MANET

There are two basic types of network layer attacks in MANET namely active and passive attacks, which are further classified into various types. Difference between them is shown in Table 2 [1].

TABLE 2 ACTIVE VS. PASSIVE ATTACKS

Active Attacks	Passive Attacks
In this, attacker disturbs the operation of the Network.	In this, attacker does not disturb the operation of the Network.
Attacker‘s goal is to modify or harm the system.	Attacker‘s goal is just to obtain information not to modify or harm the system.
Active attacks are easy to identify because the network operation is irregular.	Passive attacks are very difficult to identify because the network operation is regular.
Examples: Black hole, Gray hole, Sybil, Sleep deprivation, Rushing, Malicious Packet Dropping.	Examples: Eavesdropping, Traffic Analysis, Location Disclosure.

This research work focuses on two active attacks namely Blackhole and Malicious Packet dropping.

A. *Blackhole Attack*

In this attack a malicious node utilizes the routing protocol to claim itself of being the shortest path to the destination node. When path is selected by the routing protocol, it starts dropping the routing packets and does not forward packets to its neighbors [4]. The way the intruder node initiates the Blackhole attack may vary in different routing protocols. In case of AODV Protocol, the destination sequence number (DSN) and hop count is used to perform this attack. Destination sequence number is used to represent the freshness of the route. A high value of destination sequence number means a fresher route. To convince the target nodes, the attackers may reduce the hop count data or increase the destination sequence number (DSN).In addition, the attackers can also combine both techniques to increase severity of attacks.

As shown in the Fig. 2, an attacker node M, listens to communication channel of node S. Node M sends a forged RREP to node S immediately after node S broadcasts RREQ. Using the forged RREP, node M claims that it has both valid routing path and shortest distance to the destination node D. Because node S has no knowledge about node D in previous, node S will consider the message from the attacker as legitimate route message. Complete mechanism of this process is shown in Table 3 to Table 6. Route Request (RREQ) and Route Reply (RREP) messages exchange in route discovery process are shown in table 3 and table 4. Node S will update its routing table as indicated in table 5. Due to this attack, node S also rejects the legitimate RREP from node B [5].

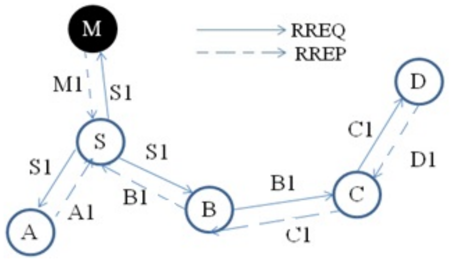


Fig. 2 BlackHole Attack

TABLE 3 ROUTE REQUEST MESSAGES

LastHop	S	S	S	A	B	C
Next Hop	M	A	B	S	C	D
RREQ	S1	S1	S1	--	B1	C1
HopCount	0	0	0	1	1	2
DSN	1	1	1	1	1	1
Origin	S	S	S	S	S	S
Dest	D	D	D	D	D	D

TABLE 4 ROUTE REPLY MESSAGES

LastHop	M	D	C	B
Next Hop	S	C	B	S
RREP	M1	D1	C1	B1
HopCount	1	0	1	2
DSN	1	1	1	1
Origin	S	S	S	S
Dest	D	D	D	D

TABLE 5 ROUTING TABLE UNDER BLACKHOLE ATTACK

S Routing Table Under Attack			
Destination	**NextHop**	**DSN**	**HopCount**
S	0	0	0
D	M	1	2

TABLE 6 ROUTING TABLE WITHOUT BLACKHOLE ATTACK

S Routing Table without Attack			
Destination	NextHop	DSN	HopCount
S	0	0	0
D	B	1	3

In Network simulator, any node is selected as Blackhole through TCL (Tool command Language) script. In AODV.cc file, the following changes is to be done in Receive_Route_Request function to perform Blackhole attack.

```
1. If (Node=Blackhole){
2. Generate Max_Sequence_Number
3. Send Route_Reply with
   Max_Sequence_Number and Hop_count=1
4. }
```

B. Malicious Packet Dropping

Once the path is established between source and destination nodes, the source node starts sending the data packets to next nodes in the path and so on. All routing protocols in MANETs are generally based on the assumption that all the participating nodes are fully cooperative. But some nodes may participate in the route discovery and maintenance processes but refuse to forward data packets. Such nodes are called as selfish or misbehaving nodes. This is also known as a data packet dropping attack. Packet dropping attacks differ from blackhole attack because there is no attempt to capture the routes in the network. To perform this attack the following changes is to be done in Route_Resolve function of AODV.cc file:

1. If (Node=Selfish)
2. Drop Packet

Selfish node is selected through TCL script. While routing data packets, Route_Resolve function is used to select next hop node.

IV. Proposed Solution to Prevent Blackhole Attack

In case of AODV Protocol, when a node receives a packet of type REPLY, it calls AODV: recvReply (Packet *p) function. Node updates its route table entry if sequence number of incoming packet is greater than previously stored number in routing table or Hop count is less than previously stored value in routing table. Proposed Algorithm to prevent Blackhole Attack is as follows:

```
1. If(SSN < DSN) or
2. (SSN = DSN) and
3. (SHC>CHC){
4. if (SSN < DSN) and (CHC!=1){
5. Call procedure Update Route Table entry}
6. if (SSN < DSN) and (CHC==1){
7. if (Source=Destination){
8. Call procedure Update Route Table entry}}}
```

Where

SSN=Stored Sequence Number

DSN=Destination Sequence Number

SHC=Stored Hop Count

CHC=Current Hop Count

Now if Blackhole Node send Route Reply with Hop count=1, Packet is rejected and no updation is done in routing table. But in case, if genuine nodes send reply with hop count = 1, Routing table is updated. If(source=destination) in line 7 of algorithm means if destination IP address send by source in Route Request Packet is equal to Destination IP address send by destination node in Route Reply Packet. It means that there is no intermediate node between source and destination. So Route Reply by any node with hop count equal to one is always rejected except if destination IP address by source in RREQ is equal to destination IP address by destination node in RREP [10]. Flowchart for this process is shown in Fig 3.

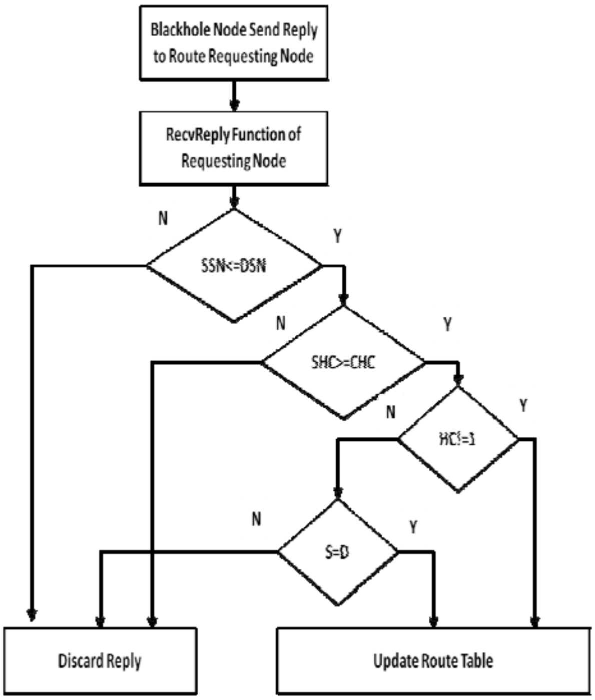

Blackhole Node Send Reply to Route Requesting Node
RecvReply Function of Requesting Node
N
Y
SSN<=DSN
N
Y
SHC>=CHC
N
Y
HC!=1
N
Y
S=D
Discard Reply
Update Route Table

Fig. 3 Flowchart for Proposed Algorithm to Prevent Blackhole Attack

V. Preventing Malicious Packet Dropping Attack

The watchdog method maintains a buffer that contains recently sent packets. This helps in detecting misbehaving nodes. When any node forwards a packet, it is ensured by node's watchdog that the next node in the path also forwards the packet. If the next node does not forward the packet then it is termed as misbehaving. This is done by node's watchdog by listening to all nodes promiscuously. In this scheme, every packet overheard by the watchdog is compared with the packet in the buffer. A match confirms successful delivery of the packet and it is removed from the buffer. Beyond the timeout period, if a packet remained in the buffer then a failure counter for the node responsible for forwarding the packet is incremented. Node is termed as malicious if this counter exceeds a predetermined threshold. Network is informed by the node that detects the problem sending a message. Flowchart for this process is shown in Fig. 4 [7].

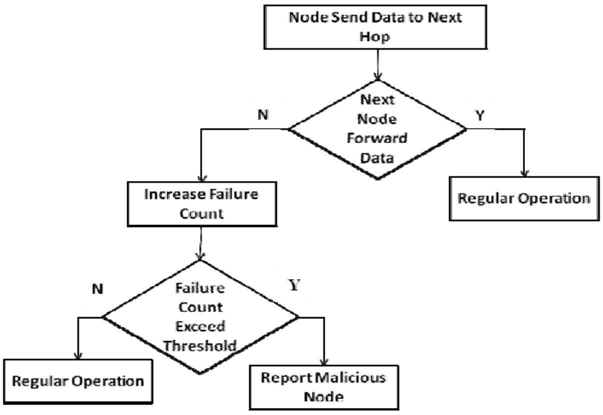


Fig. 4 Flowchart for Watchdog Intrusion Detection System

VI. Work Flow

In this Research Work, Blackhole and Malicious Packet Dropping attack is simultaneously performed on AODV protocol. Proposed solution to prevent Blackhole attack is combined with Watchdog IDS. Main aim of this simulation is

- To analyse performance of AODV protocol in presence of dual attack (i.e. Blackhole and Malicious Packet Dropping) without any prevention scheme.
- To analyse performance of AODV protocol in presence of dual attack with prevention scheme.

Performance metrics calculated under these scenarios are discussed in section VII.

VII. Simulation Environment and Performance Metrics

The simulation is conducted within the Network Simulator (NS) 2.35 environment on a Ubuntu 12.10 operating system. The system is running on a laptop with Core i3 CPU and 4-GB RAM. Each network scenario is run five times and Average is calculated.

The various parameters analyzed and measured are as follows:

A. *Average end-to-end delay (AED)*

It is calculated for each data packet by subtracting the sending time from the received time of the packet at final destination.

$$AED = \frac{\sum_{1}^{N} (T_R - T_S)}{N}$$

Where

N = Number of successfully received packets
T_R = Packet Received Time
T_S = Packet Sent Time

TABLE 7 Simulation Parameters

Channel type	Wireless channel
Number of nodes	100
Traffic type	CBR
Data Payload	512 bytes/packet
MAC Types	802_11
Node Placement	Random
Mobility	Random way point
Node Speed	10 m/s
Area of simulation	1000m X 1000m
Number of Malicious Nodes (Selfish+Blackhole)	2–10
Time of simulation	150 sec
Protocol	AODV

B. *Packet Delivery Ratio (PDR)*

It is the ratio of the number of packets received by the destination node to the number of packets sent by the source node.

$$PDF = \frac{\sum \text{Received packets at}}{\sum \text{Sent packets by sources}}$$

C. *Routing Overhead (RO)*

It defines the ratio of the amount of routing-related transmissions (i.e. Route REQuest (RREQ), Route REPly (RREP), Route ERRor (RERR)) to total transmissions (i.e. data transmission and routing transmission).

$$RO = \frac{\sum \text{Routing transmissions}}{\sum \text{Data transmissions} + \sum \text{Routing transmissions}}$$

VIII. Result Analysis

A. *Packet Delivery Ratio*

Packet delivery ratio generally decreases in presence of Selfish and Blackhole nodes. As shown in Fig. 5, PDR decreases with increase in number of malicious nodes in case of standard AODV protocol. But decrease is less in case of Modified AODV protocol.

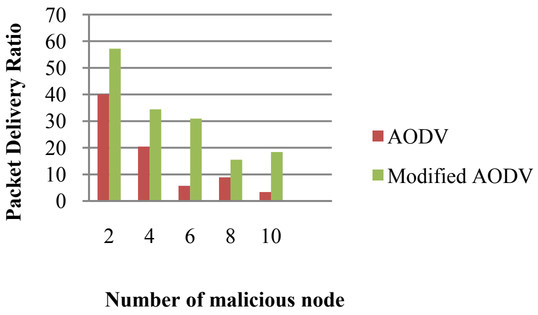


Fig. 5 Packet Delivery Ratio in Presence of Malicious Node

B. Routing Overhead

Routing overhead generally increases in presence of selfish and Blackhole nodes. As shown in fig.6, RO increase with increase in number of malicious nodes in case of standard AODV protocol. But increase is very less in modified AODV as compared to standard AODV protocol.

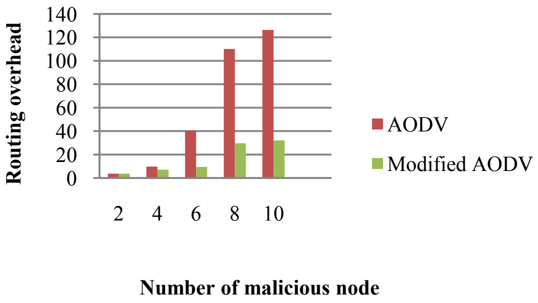


Fig. 6 Routing Overhead in Presence of Malicious Node

C. Average End-to-End Delay

Average End-to-End delay shows fluctuating behavior in both cases as shown in fig.7. Reason behind this behavior is that, nodes are randomly moving in 1000 m*1000m area. Sometimes sender and receiver nodes are close to each other and sometimes they are far apart from each other. Besides this reason, malicious nodes are also randomly selected; therefore number of malicious nodes in path increases or decreases.

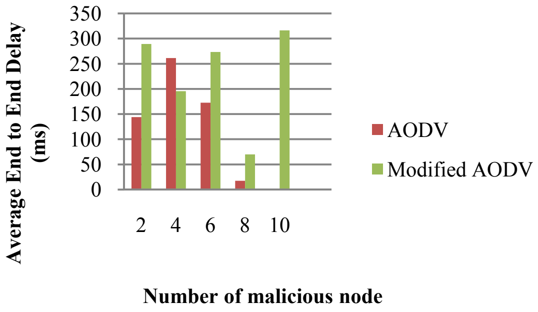


Fig. 7 Average End-to End Delay in Presence of Malicious Nodes

IX. Conclusion

The work presented here is primary concerned with security issues in Mobile Adhoc Network (MANET).The performance of AODV protocol is analyzed in presence of Blackhole and Malicious Packet dropping attack. Both these attacks are performed simultaneously on the network. Some nodes are programmed as Blackhole nodes and some as selfish. Packet delivery ratio decreases and routing overhead increases in presence of these attacks. Solution is proposed to prevent Blackhole attack. Malicious packet dropping attack is prevented by using Watchdog Intrusion detection System. Proposed solution to prevent Blackhole attack is combined with watchdog intrusion detection system in order to prevent both attacks. Modified AODV protocol shows better results i.e. increases packet delivery ratio and decreases routing overhead as compared to standard AODV Protocol.

References

[1] D. Djenouri, L. Khelladi, N. Badache “ A survey of security issues in mobile ad hoc networks”. IEEE communications surveys, Vol 3. No7, 2005.Retrived from: http://www.lsi-usthb.dz/Rapports_pdf /2004/ LSIIR-TR0504.pdf.

[2] M.Schutte, ”Detecting Selfish and Malicious Nodes in manets”. seminar: sicherheit in selbstorganisierenden netzen, hpi/ universität potsdam, sommersemester, 2006.Retrieved from: http://mschuette.name /files/uni /soN-text.pdf

[3] A.Nadeem, M.P Howarth, ”A Survey of MANET Intrusion Detection & Prevention Approaches for Network Layer Attacks”. IEEE communications surveys & tutorials.Vol 15 No 4 PP. 2027-2045.2013. doi:10.1109/surv.2013.030713.00201

[4] F.H Tseng, L.D Chou, H.C Chao.” A survey of black hole attacks in wireless mobile ad hoc networks “Human-centric Computing and Information Sciences VOL 1. NO 1, PP 1-16. Retrivedfrom:http://link.springer.com/article/10.1186/2192-1962-1-4.

[5] S.Mandala, A.H Abdullah, A.H. A.S Ismail, H.Haron, A.H Ngadi, Y.Coulibaly, .” A Review of Blackhole Attack in Mobile Adhoc Network”. 3rd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME) Bandung, November 7-8.339-344. doi:10.1109/ICICI-BME.2013.6698520.

[6] P.Sahu, S.K Bisoy, S.Sahoo, ”Detecting and Isolating Malicious Node in AODV Routing Algorithm”. International Journal of Computer Applications, vol.66 No.16, PP 8-12, 2013.

[7] J.Hortelano, ”SafeWireless”, Avaiable at http://sourceforge.net/ projects/ safewireless/files/.

[8] A.Al-Roubaie, T.Sheltami, A.Mahmoud, E.Shakshuki, H. Mouftah, ”.AACK: Adaptive Acknowledgment Intrusion Detection for MANET with Node Detection. Paper presented at 24th IEEE International Conference on Advanced Information Networking and Applications.2010. DOI 10.1109/AINA.2010.136

[9] A.M Kanthe, D. Simunic, R. Prasad. "Effects of malicious attacks in mobile ad-hoc networks." Computational Intelligence & Computing Research (ICCIC), 2012 IEEE International Conference Pp. 1-5, 2012.

[10] K.Singh, A.Kaur, ” Security Enhancement in AODV Protocol against Blackhole Attack” Paper Presented at *NCISC (National Conference on Information Security Challenges)* held in march 2014 at Babasaheb Bhimrao Ambedkar University(A central university) Lucknow, Vol 1 No 1, Pp. 59-64, 2014 .

Impact of Denial of Service Attack on the Virtualization in Cloud Computing

Kanika[1] and Navjot Sidhu[2]
[1,2]*Centre for Computer Science and Technology, Central University of Punjab, Bathinda, India*
E-mal: [1]kanikagoyal05@gmail.com, [2]navjotsidhu8@gmail.com

***Abstract*—Cloud computing is the fastest growing technology in the IT world. The technology offers reduced IT costs and provides on the demand services to the individual users as well as organizations over the Internet. Cloud computing offers infrastructure as a service (IaaS). IaaS provides infrastructure including software, hardware, storage space, network bandwidth to the users on demand over the internet. Cloud computing makes use of virtualization to provide infrastructure as a service. Virtualization is based on the concept that multiple tenants can use the same physical machine with multiple operating systems. Virtualization comprises the vulnerability of Denial of Service (DOS) attack that can affect the performance of cloud computing. A malicious VM attacker can compromise the other guest VM or the host OS. The paper explores the TCP SYN flood attack over the other guest VM by a malicious VM attacker co-existing in the virtualized cloud infrastructure. Different Parameters are analyzed over the victim VM to detect the TCP SYN flood attack**

Keywords: Cloud Computing, Virtualization, Hypervisor, Network Security

I. Introduction

Cloud computing is the Internet-based computing, where sharing of resources, software and platforms are provided to the users on demand in a distributed computing environment. Cloud computing is the growing trend for storing and processing data in a resource sharing environment. The term cloud in the cloud computing specifies storage space, hardware, networks combination to deliver computing services. Cloud services include delivery of software, platform to develop applications and providing a complete infrastructure over the Internet. Cloud computing relies on sharing of computing resources rather than having local servers. Cloud computing creates exciting opportunities like reduced costs and flexibility to the users.

A. *Cloud Computing Service Models*

Cloud service providers offer services that are separated into three categories as [1]:

1) *Software as a Service (SaaS)*

In SaaS, software are offered as a service on demand to the users. Users are billed on the basis of usage and there is no need for investment in servers or software licenses.

2) *Platform as a Service (Paas)*

PaaS provides complete platform required to develop user specific applications and services over the Internet. Platform as a service offers combination of operating system and application servers, such as Linux, Apache, MySql and PHP etc.

3) *Infrastructure as a Service (IaaS)*

IaaS offers complete infrastructure such as servers, basic storage systems, networking equipments over the Internet. Here multiple tenants share a virtualized environment. Tenants are coupled with managed services for OS and application support.

B. *Essential Characteristics*

The five characteristics of the cloud which represents its services are [10] [12]:

1) *On-demand self-service*

Consumers can automatic provision computing resources without requiring interaction with cloud service provider.

2) *Broad network access*

Cloud services are provisioned over the network and can be accessed via multiple devices such as mobile phones, laptops, PDA, etc.

3) *Resource pooling*

The cloud service provider's resources are pooled in a multi tenant environment. Resources are dynamically allocated to the tenants according to their demand. The tenants don't know the exact location of the resources. The shared resources include storage, processing, memory, etc.

4) *Rapid elasticity*

Cloud services can be automatically scaled at any time and at any quantity depending upon the user's demand.

5) *Measured service*

Customer's usage of the provider's services is automatically monitored and reported providing transparency for both the customer and provider.

II. Multi-Tenancy and Virtualization

In a multi-tenant environment, tenants have their own private space to save private data as well as global space shared among all tenants. By sharing resources and creating standard offerings, multi-tenancy offers reduced cost and optimum use of resources in a shared environment [1].

With SaaS, data of multiple tenants is stored on the same database and may share the some tables. In IaaS, multiple tenants share infrastructure resources such as hardware, servers and storage devices [4] [13].

Resources shared among multiple tenants can be:

1. Basic storage space.
2. CPU processing.
3. Memory.
4. Network bandwidth.

Multi-tenancy is obtained by the use of virtualization. It allows multiple operating systems to run on a single machine simultaneously. In cloud computing virtualization used to serve several end users by creating virtual version of storage space, operating system, hardware platform [16].

Virtualization divides a physical computer to several virtual machines known as guest machines. Multiple virtual machines run on a host computer, each having its own OS and applications. Virtualization gives an illusion to the users that they are running their processes on a physical computer independently, but in reality they are sharing the resources of a single host machine. The software which permits multiple operating systems to use the resources a physical machine is called a hypervisor. The hypervisor resides between the operating system of the host machine and the virtual environment [4] [14].

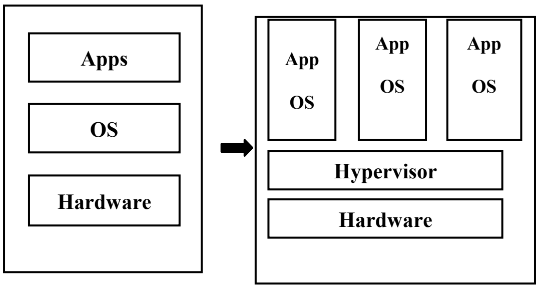


Fig. 1 Independent OS to Virtualization of OSs

The Fig. 1 shows how an individual operating system running its applications on the independent physical hardware can be placed in a virtual machine.

All the OSs share the same physical system with other virtual machines. The machine with administrative capabilities lower to hypervisor is said to be Host machine which controls the hypervisor and other virtual machines said to be guest OS.

As the tenants sharing the same physical host with unknowns, there are various kinds of risks associated with the multi-tenancy environment in the cloud computing. So it is important aspect to isolate the multiple users on same physical [5], [6].

The hypervisor, a software layer which manages the virtualization, allows virtual machines to execute simultaneously on a single machine. This provides hardware abstraction to the running Guest OSs and efficiently manages underlying hardware resources. There are numerous hypervisors ranging from open-source such as KVM, Xen and virtual box, to commercial hypervisors such as VMware vSphere and Microsoft Hyper-V etc [11].

III. Security in Multi-Tenant Environment

As the multiple tenants sharing the same physical host with unknowns, there are various kinds of risks associated with the multi-tenancy environment in the cloud computing. An attacker may use guest OS (Virtual Machine) try to communicate and compromise other Virtual Machines on the same physical host, therefore breaking the isolation characteristic of VMs. The most common attacks under this are Measure cache usage, Sniffing attack, Spoofing attack, denial of Service (DoS) attack [7], [13] .

A. *TCP DDOS Attack*

In virtualization environment, resources such as CPU, memory, disk and network are shared by VMs and the host. An attacker aims to exhaust the resources from a physical host in order to deny service to the other VMs in the machine [8], [15].

Denial of service attack [2], [3] is one of the most dangerous network attacks, in which the one victim machine receives more TCP-SYN requests than its capacity, so that other machines requests could not be served by the main host in the cloud environment. TCP SYN Flood attack can be most dangerous than unclouded environment because of VMs are sharing their resources with the neighbour VM and Host. Under TCP SYN Flood, one virtual machine is used as a source of denial of service attack to another virtual machine present in same infrastructure.

IV. Related Work

TCP is a connection oriented protocol that needs “handshaking” to start communication in client-server architecture. The protocol provides reliable delivery of data. The client sends a “SYN” packet to server to whom it wants to establish the connection.

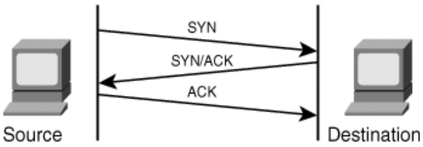


Fig. 2 TCP Three Way Handshake

The server replies with a “SYN/ACK” packet that to accept the connection. Then the client sends an

"ACK" packet to establish the connection. The connection complete connection is established in three steps, so the procedure known as "Three Way Handshaking" [2].

A. *TCP SYN Flood*

TCP 3-way handshake structure is exploited to perform Denial of service attacks by TCP SYN flood. The attacker overloads the victim with large number of TCP connection requests and it will not able to respond to legitimate requests.

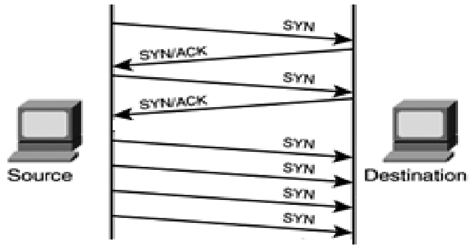


Fig. 3 TCP SYN Flood

The victim saves each new TCP connection to its buffer and transmits SYN-ACK packet to establish the connection. The attacker does not respond to the SYN-ACK. A large number of half open connections are left on the victim's queue and it gets overflow. The queue of the server is limited, and legitimate client's request cannot be fulfil due to unavailability of the resources (space) in the queue [3] [9].

B. *IP Spoofing*

IP spoofing is done by the attacker to create the IP packets with forged IP source address. In DoS attack, the attacker uses the IP spoofing to flood the TCP SYN packets from false IP identity. The attacker does not care about receiving response back to the IP packet. IP spoofing uses randomized IP addresses to start the three way handshake. Spoofed IP addresses are difficult to filter since each spoofed packet appears to come from a different address. The attackers also use subnet spoofing, spoofs a random address within the address space of the sub network [17].

V. Experiment Architecture

To conduct the experiment, the private cloud infrastructure is deployed using VMware ESXi and vSphere client. The physical server VMware ESXi hosted hypervisor is installed that provides sharing of different resources such as the CPU, memory, Network Interface Card (NIC) to multiple VMs. The vSphere Client is the interface that accesses and manages the multiple the VMs remotely.

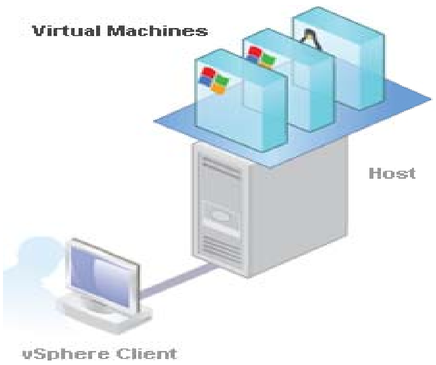


Fig. 4 Virtualized Cloud Infrastructure

Ten guest virtual machines are installed over the hypervisor and accessed through the vSphere client. Among the guest OS (VMs) one machine with the IP address 192.168.43.129 is the malicious node and sniffs the network traffic to know about the other tenants present in the network. The attacker VM acts as a source of the TCP SYN flood packets to another VM existing in the same network. The victim VM with the IP address 192.168.43.138 and TCP backlog 1024, receives TCP SYN packets more than its capacity, and its resources get exhausted. The other virtual machines are used as Zombie that is connected on the same network segment as the host and guest virtual machines.

VI. TCP-SYN Flood Attack

Using the 'nmap' tool the attacker virtual machine performs the scan to know about the other virtual machines IP addresses present in the network.

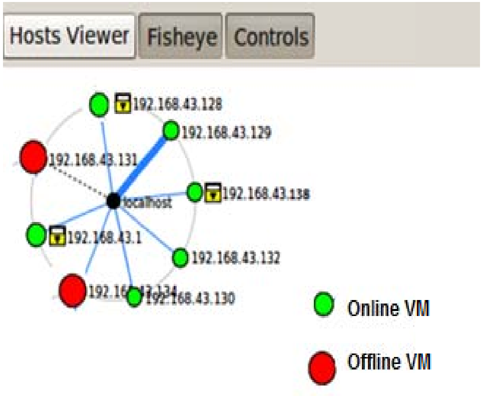
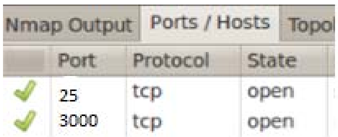


Fig. 5 Nmap Scanning Result

The VMs with green symbol are currently online, and the VMs with red symbol are currently offline in the network. The attacker VM picks the online co-existing VM with IP address 192.168.43.138 to perform TCP-SYN flood. The attacker VM scans the VM to check for the open TCP ports to perform the attack with the 'nmap'.



	Port	Protocol	State
✓	25	tcp	open
✓	3000	tcp	open

Fig. 6 ‘nmap’ Tool Scanning for the Open Ports

The scan showed for the IP address ‘192.168.43.138’ TCP port 25 and TCP port 3000 are open.

The attacker virtual machine makes use of hping3 tool to SYN flood the TCP port 3000 in a distributed manner with the direct IP address and spoofed IP addresses of other virtual machines that are offline or online in the network.

A. *Direct Attack*

The attacker VM rapidly sends TCP SYN packets with its own IP address as the source.
The command used to flood TCP SYN request is:

Sudo hping3 -flood -S -p 3000 192.168.43.138

The attacker VM initiates the TCP connection by sending SYN packets and the victim VM replies with the SYN-ACK packet, and then the attacker doesn't send the final acknowledgement to complete the three-way handshake. At the victim VM site, high numbers of half-opened connections are left.

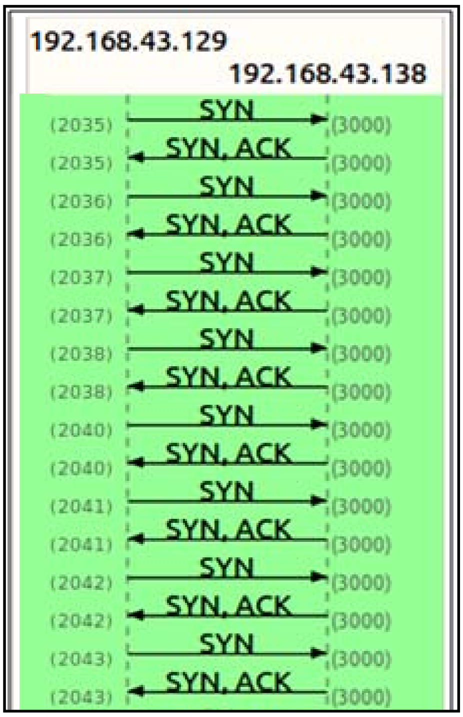


Fig. 7 SYN Flood with own IP Address

The queue that is storing the half opened connections is of finite size that can have 1024 backlog at any instant of time, and it is made to overflow by intentionally creating too many half-open connections. The victim keeps on waiting for the final ACK packet and after the RTT (round trip time) expires; it resends the SYN-ACK packets to the attacker. The victim VM is not able to further create new TCP sessions for the legitimate network traffic.

B. *IP Spoofing with Offline VM*

The attacker floods the TCP SYN packets with the spoofed IP addresses of other co-existing VMs that are offline at that instant.
The command used to flood TCP SYN request is:

Sudo hping3 -flood -S -p 3000 192.168.43.138 – a 192.168.43.131

Sudo hping3 -flood -S -p 3000 192.168.43.138 – a 192.168.43.134

In a short period of time there are a number of connection attempts by the IP 192.168.43.131 and 192.168.43.134 to the VM 192.168.43.138. Within a private network When the VM wants to send data to the co–existing VM, ARP cache is used to find out the MAC address corresponding to the VM.

The victim VM tried to resolve the MAC address of the VMs (offline). But when no response is received by the offline VMs, the victim VM not having the physical address of the host, it cannot send an ACK-SYN to the same to continue with the three-way handshaking. The TCP/IP stack of the server has to wait for a set time for each connection. During this time more packets keep arriving that create new connections. At the victim side, for each connection that tries to be made, a structure in memory called TCB (Transmission Control Block) is created.

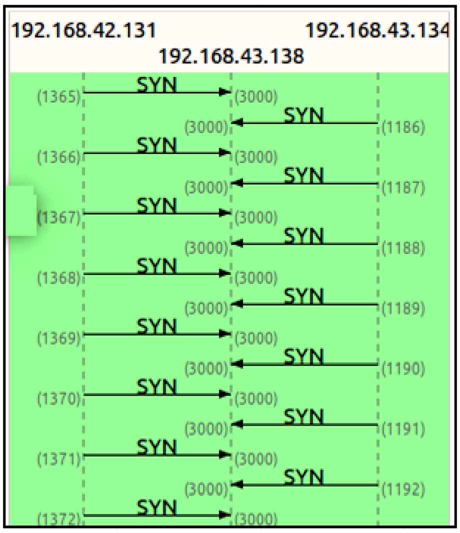


Fig. 8 SYN Flood with Offline VM Spoofing

The TCB holds the SYN packet information before the connection is fully established. It holds only 1024 half opened connections. The attacker sends SYNs that causes the allocation of so many TCBs that a victim VM’s kernel memory is exhausted.

C. IP Spoofing with Online VM

The attacker VM sends SYN packets to the victim VM, with the spoofed IP addresses of the VM that are online on the same network. The spoofed VMs act as zombie.

The command used to flood TCP SYN request is:

Sudo hping3 -flood -S -p 3000 192.168.43.138 – a 192.168.43.130

Sudo hping3 -flood -S -p 3000 192.168.43.138 – a 192.168.43.132

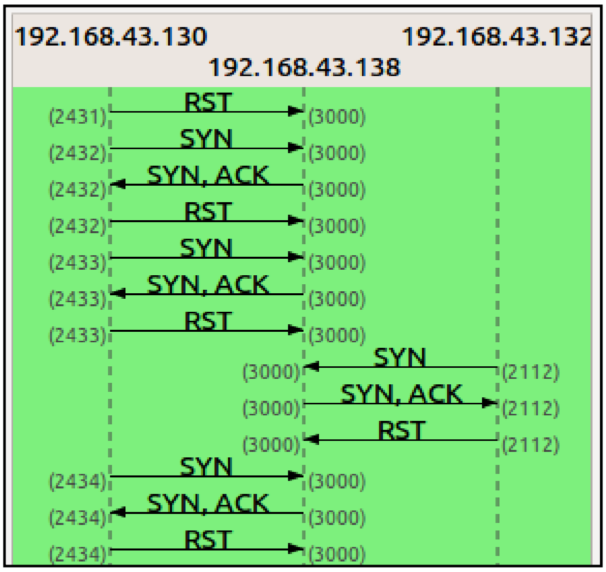


Fig. 9 SYN Flood with Online VM Spoofing

As the flow graph shows the victim VM sends the ACK-SYN packets to the respective IP addresses. The zombie VM won't be expecting the SYN/ACK (because, it has not really sent the SYN), so the zombie VM responds to the victim VM with a RST. The attacker keeps the victim busy in handling the spoofed packets and consuming the resources. The victim VM's resources are depleted; it is not further create new TCP sessions legitimate network traffic.

VI. Detection of TCP-SYN Flood

To detect the attack effect, the attacker Virtual Machine trying to communicate with the victim Machine. 20 seconds after communication, attacker starts sending attack traffic that lasts for 40 seconds. The attacker virtual machine floods the victim at the maximum possible rate allowed by operating system.

Wireshark, Bandwidth monitor, Netflow, Netstat commands and IPtraf are few of the tools used to analyze the system under attack .The research to measure the performance of victim virtual machine over the TCP DOS attack by a malicious guest VM. The performance of the victim VM under attack is determined on the basis of network traffic, average number of SYN requests over the system, SYN to FIN|RST ratio, resource utilization, etc.

A. Number of SYN Requests Captured

The SYN packet is sent to initiate the TCP Three-way handshake. The attacker floods the victim VM by sending a large number of TCP SYN requests. Wireshark captures the SYN packet passing through the eth0 port. The Ethernet port was monitored during a TCP SYN flood attack.

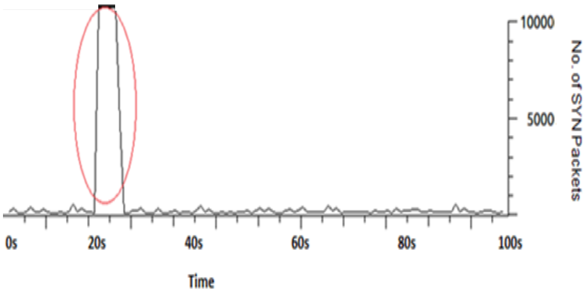


Fig. 10 Number of SYN Packets at the Victim VM with Attack

The Fig. 10 shows the result of the incoming traffic for the TCP Port 3000. During TCP-SYN flood attack (from 20 sec to 30 sec) the number of SYN requests more than 10000 as compared to normal traffic that is about 5 to 10 SYN requests per second.

B. SYN and FIN/RST Packet Ratio

TCP is a bi-directional protocol. The TCP connection is terminated by the FIN packet. The TCP connection performs half-duplex termination by sending RST packet from either side. The RST packet aborts the TCP connection. The number of FIN packets and the SYN packets are almost same under the normal TCP sessions. TCP session may be terminated by a RST packet without a FIN packet. But when the attack occurs, the relation between the SYN packets and FIN|RST Packets completely breaks. Detection of TCP SYN Flood is done based on the change of the difference between the number of SYN and the number of RST | FIN.

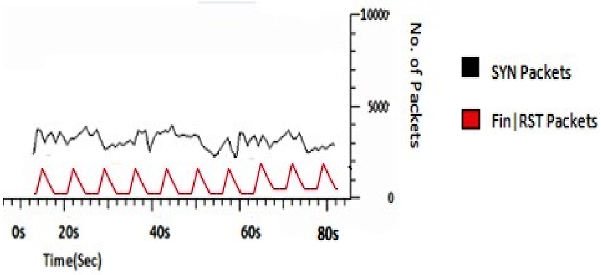


Fig. 11 Normal SYN to FIN|RST Packet rate

The Fig. 11 shows that the number of SYN and FIN|RST packets is almost same under normal network behaviour. The number of connections opened by the legitimate users is equal to the number of connections closed under the normal TCP session.

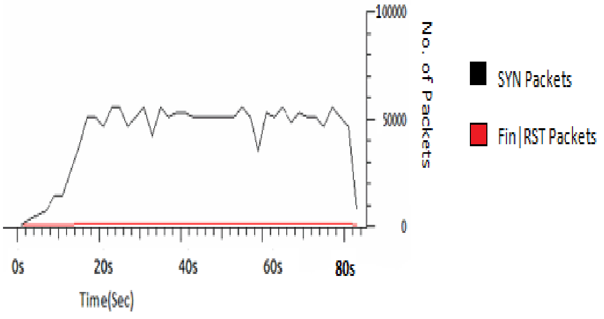


Fig. 12 SYN to FIN|RST Packet Rate with SYN Flood

When the attacker performs the SYN Flooding to the VM, it doesn't terminate the connection at the victim VM side. The Fig. 12 shows the number of SYN and FIN|RST packets rate when the system is under attack. The number of SYN requests is very high as compared to the FIN|RST packet which is almost zero.

C. The Start and End Time of an Attack

The exact time when the attack starts is analyzed with the post processing of the TCP SYN packets. Incoming traffic rate increases abruptly during the TCP SYN flood attack as compared to normal traffic rates.

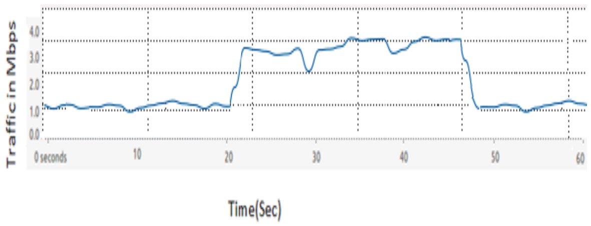


Fig. 13 Time duration of attack

From the Fig.13, it could be seen that the normal incoming traffic rate is almost 1 Mbps and the traffic rate goes up to 3Mbps at the time of TCP SYN flood from 20thsec to 50th sec. The SYN Flood attack is detected based on the incoming traffic rate that increases abruptly as compared to the traffic rates under normal network behavior.

D. Resource Utilization on the Host OS

As under the virtualized cloud infrastructure the single CPU is shared among multiple VMs. CPU utilization refers to hypervisor's usage of processing resources. For each TCP connection, that tries to be established, a queue is maintained in the memory that holds all the information about a TCP connection.

It could be seen from the figure that CPU % utilization for the single virtual machine increases to 65% when it is under the attack. The memory utilized by the victim VM under the attack is up to 30%.

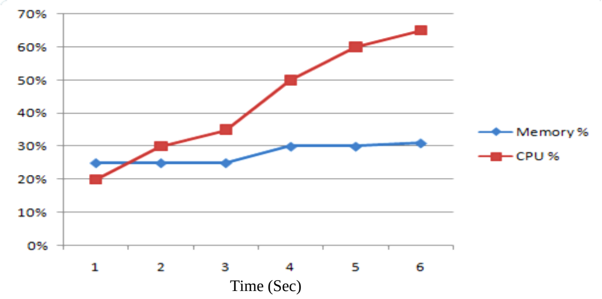


Fig. 14 Resource Utilization of Host OS

VII. CONCLUSION

Multi-tenancy in virtualization not only allows more effectiveness of the infrastructures to the cloud service providers, but also introduces new attack vectors in the cloud. Cloud computing security issues need to be approached cautiously. The paper includes the experiment that shows the vulnerability that how a malicious virtual machine can attack over another virtual machine in a virtualized cloud. The vulnerability of the Denial of Service attack by a malicious virtual machine over co-existing virtual machine in the private cloud infrastructure is explored along with a mechanism on how to approach it. The malicious virtual machine exhausted the common resources by flooding the co-existing VM with high rate of unreasonable network traffic.

The malicious virtual machine is detected on the basis of different parameters over the victim operating system. Network Traffic is analysed over the victim VM. The traffic over the victim increases at a very high rate as compared to average traffic whenever there is an attack in the system and corresponding to that more resources wastages at the victim. The results showed that the arrival rates of normal TCP SYN packets and attacked SYN Flood varies with large difference. On the basis of daily network behaviour a SYN Packet arrival rate is decided. The presence of TCP-SYN Flood attack is determined based on the average number of SYN requests to the VM, SYN to FIN|RST packet ratio. This research may prove to strengthen virtualization and reduces the risks of cloud computing. Immediate extensions to the research work include prevention and mitigation of TCP SYN Flood by configuring the firewalls at the VM level and the hypervisor level.

REFERENCES

[1] A. Jasti, P. Shah, R. Nagaraj, R. Pendse "Security in multi-tenancy cloud, " in IEEE International Carnahan Conference on Security Technology (ICCST), pp.35-41, 2010.

[2] A.Habib, M. Hefeeda, B. Bhargava, "Detecting service violations and DoS attacks" 2003.

[3] A. Bakshi, B. Yogesh, "Securing cloud from DDOS Attacks using Intrusion Detection System in Virtual Machine, " in Second International Conference on Communication Software and Networks, pp. 260-264, 2010.

[4] B. Grobauer, T. Walloschek, E. Stocker, “Understanding Cloud Computing Vulnerabilities, ” Security & Privacy, IEEE, vol. 9, pp. 50-57, 2011.

[5] CSA, “Security Guidance for Critical Areas of Focus in Cloud Computing V2.1” Cloud Security Alliance, 2009, Available: https://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf [Accessed: 08-Jan-2014].

[6] G. Wang, T.S.E. Ng, “The impact of virtualization on network performance of amazon ec2 data center, ” in Proc. IEEE INFOCOM, pp. 1–9, 2010.

[7] H. Wu, Y. Ding, C. Winer, L. Yao, “Network security for virtual machine in cloud computing, ” in Proc. 5th International Conference on Computer Sciences and Convergence Information Technology, pp.18-21, 2010.

[8] M. A. Bamiah, S. N. Brohi, “Seven Deadly Threats and Vulnerabilities in Cloud Computing, ” International Journal Of Advanced Engineering Sciences And Technologies, Vol No. 9, Issue No. 1, pp. 87 – 90, 2011.

[9] N. H. Bhandari, “Survey on DDoS Attacks and its Detection & Defence Approaches, ” International Journal of Science and Modern Engineering (IJISME), pp. 67-71, 2013.

[10] P. Mell, T.Grance, “The NIST definition of Cloud Computing, ” NIST, Special Publication 800–145, 2011.

[11] P. Nomnga, M. S. Nyambi Scott, “Technical Cost Effective Network-Domain Hosting through Virtualization: a VMware ESXi and vSphere Client Approach, ” International Journal of Computer Applications. Pp. 39-47, 2014.

[12] R. Buyya, J. Broberg, A. M. Goscinski, “Cloud Computing: Principles and Paradigms, ” vol. 87, John Wiley & Sons, 2010.

[13] S. Subashini, V. Kavitha, “A survey on security issues in service delivery models of cloud computing, ” Journal of Network and Computer Applications, pp. 1-11, 2011.

[14] S. Brohi, M.Bamiah, “Challenges and Benefits for Adopting the Paradigm of Cloud Computing”, International Journal of Advanced Engineering Sciences and Technologies (IJAEST), vol. 8, pp. 286 - 290, 2011.

[15] S. N. Brohi, “Identifying and analyzing security threats to Virtualized Cloud Computing Infrastructures, ” in International Conference on Cloud Computing Technologies, Applications and Management (ICCCTAM), pp. 151-155, 2012.

[16] W. Dawoud, I. Takouna, C. Meinel, “Infrastructure as a service security: Challenges and solutions, ” In the 7th International Conference on Informatics and Systems (INFOS), pp. 1-8, 2011.

[17] Y.S.Choi “Integrated DDoS attack defense infrastructure for effective attack prevention, ” in IEEE International Conference on Information Technology Convergence and Services, pp.1-6, 2011.

A Survey on Distributed Localization Techniques in Acoustic Underwater Sensor Network

Manpreet[1] and Jyoteesh Malhotra[2]
[1,2]*Department of Computer Science & Engineering,*
Guru Nanak Dev University Regional Campus, Jalandhar, India
E-mail: [1]manpreetpawar17@gmail.com, [2]jyoteesh@gmail.com

Abstract**—Underwater acoustic sensor network is an enabling technology for detecting the animals in water, other aquatic applications and at such situations where human operation is not possible under water. It consists of sensor nodes and vehicles deployed underwater to perform monitoring task. Recently it gained much attention. Because of unapproachable underwater situation, the task of finding a location of underwater sensor node will become pivotal. In contrast terrestrial localization, it is complicated. Various localization techniques for underwater acoustic sensor network have been proposed in the literature. This paper has thoroughly surveyed these technique, their issues and challenges have been extracted which will act as a guideline for active researches.**

Keywords: Underwater, Localization, AUV, UW-ASN

I. Introduction

Underwater acoustic sensor network (UW-ASN) is taking important place because it is needed to monitor the aquatic environment. It is useful in many applications as for underwater environment monitoring, biomedical health monitoring, detecting the river or sea pollution, oil monitoring, disaster prevention and oceanographic data collection needed to protect the aquatic life. Localization is one of the major and challenging tasks in UW-ASNs used for track the nodes, improve the medium access control and network protocols. UW-ASN consist Multiple unmanned or autonomous underwater vehicles (UUVs, AUVs), equipped with under water sensors to collect the data. To achieve this objective, sensors and vehicles self-organize in an autonomous network; all these can adapt the properties of ocean [1]. There are some vehicles exists which are used for aquatic purposes. These vehicles are used in UW-ASN localization. Localization in UW-ASN is different than terrestrial localization because of GPS un-availability under water, radio frequency signals have limited propagation and UW-ASN faced equipment failure due to corrosion attack and other attacks faced under water so terrestrial localization techniques can not apply for UW-ASN localization. Localization techniques for UW-ASNs should tackle lack of GPS, 3-D space, mobility, minimal message exchange and try to be robust under sleep/wake-up cycles [3]. Architecture of USN designed carefully so that it can efficiently adapt the UW-ASN. Architecture should be optimized whenever possible [1]. There are following types of architectures in UW-ASN:

A. Static two-dimensional UW-ASNs

Sensor nodes are anchored to the bottom of the water with the deep ocean anchors. It is applicable for environmental monitoring.

B. Static three-dimensional UW-ASNs

Sensor nodes float at different depth. Sensor nodes make a network. It is useful for monitoring pollution and other aquatic environment activities.

C. Three-dimensional Networks of Autonomous Underwater Vehicles (AUVs)

This architecture consist two parts. One part is fixed, consists anchored sensor nodes and mobility part contains AUV which used to broadcast the message.

The rest of the paper is organized as follows. Section II describes requirements of localization in UW-ASN. Section III describes the proposals of UW-ASN localization techniques. Section IV describes open issues and future scope before the paper is concluded in section V.

II. Requirements of UW-ASN Localization

There are many terrestrial localization techniques exist. Then why UW-ASN localization required [1]. Firstly this section will describe the difference between terrestrial and UW-ASN localization:

A. Deployment

Deployment in terrestrial is very dense but underwater it is deemed to be sparser due to cost involved and other challenge factors.

B. Power

Power needed for UW-ASN is more than terrestrial localization because the equipments consume more energy and also due to channel impairments more energy needed for localization.

C. Memory

Terrestrial sensor nodes require less memory, but UW-ASN requires cache to collect the data for localization of sensor nodes stack.

D. *Communication*

In terrestrial sensor nodes communication done through electromagnetic waves but in UW-ASN for communication acoustic channels are deployed.

E. *Cost*

Terrestrial localization equipments are not as much expensive as UW-ASN. This is because UW-ASN is very complex and there is extra protection needed for underwater equipments to protect from corrosion and other oceanic attacks.

Considering above differences of terrestrial sensor network and UW-ASN, the terrestrial localization techniques are not suitable for UW-ASN. So UW-ASN needs different localization techniques.

III. Proposals of Localization Techniques

Considering the importance UW-ASN localization various researchers have suggested the localization techniques UW-ASN as mentioned below.

Melika Erol, Luiz F.M. Vieira, and Mario Gerla [2] proposed Dive ‘N’ Rise (DNR) positioning. In this technique DNR beacons used. DNR beacons float above water and get coordinates from GPS then after diving in water these broadcast its coordinates to sensor nodes which are underwater. Simple apparatus used which can dive with weight and it releases its weight into the depth of water and rise with bladder. DNR can be used with current or without current. Without current sensor nodes do not drift and only receive and process message but with current sensor node can drift.

Melika Erol, Luiz F.M. Vieira, and Mario Gerla [3] propose AUV aided localization. In this AUV receives GPS signals while floating above water then it dives and move among sensor nodes. While moving it broadcast its message among sensor nodes. It is mainly used where sensor nodes are freely scattered in water and they have no network connection. Sensor nodes are not tied with any fixed object and can freely move.

Hanjiang Luo, Yiyang Zhao, Zhongwen Guo, Siyuan Liu, Pengpeng Chen, and Lionel M. Ni [4] propose localization using directional beacons (UDB). This technique replaces the Omni-directional localization. Omni-directional localization provides more extensive coverage but directional localization is applied for some special cases [5]. In this technique there is no need of communication between sensor node and AUV. AUV broadcast message, sensor nodes only listen the AUV. This technique is very useful in strap area and reduces energy consumption of sensor nodes. UDB can provide accurate localization.

Zhong Zhou, Jun-Hong Cui, and Shengli Zhou [6] propose localization in large scale underwater sensor network. This technique consist three types of nodes: 1) surface buoys which can drift on the water surface and equipped with GPS. These nodes can get their absolute location from GPS or by some other manner.2) Anchor nodes can communicate with surface buoys and ordinary nodes. Anchor nodes get their position from surface buoys and assist ordinary nodes to do localization. 3) Ordinary nodes which can not directly communicate with surface buoys but can communicate with anchor nodes to get their location. In [7] author shows that this technique has higher energy consumption as compared to DNR localization.

Kai Chen, Yi Zhou, and Jianhua He [8] propose Detachable Elevator Transceivers localization scheme. This technique consist three types of nodes: 1) Surface buoys which can drift on the water surface and equipped with GPS. These nodes can get their absolute location from GPS or by some other manner.2) DET which are attached to surface buoys and can dive and rise in water to broadcast its position. 3) Anchor nodes can communicate with DET. If any anchor node receive message from three DET then it calculate its location and help to locate the ordinary nodes. 4) Ordinary nodes which only listen to anchor nodes. If any ordinary node receive message from three anchor nodes then it calculate its location. This technique is also employed in [9]. This technique increase localization ratio, scalability and decrease cost of the system.

A.K. Othman, A.E. Adams, and C.C. Tsimenidis [10] propose a node discovery and network discovery localization technique. This technique is named as anchor free and also employed in [11]. In this technique one node is selected as seed node and broadcast its message to gain information from its neighbors. Then on the basis of reply from neighbors its select farthest node as seed node. Seed node repeats the process. Firstly nodes define local coordination system then network coordination system.

Xizuhen Cheng, Haining Shu, Qilian Liang, and D. H.-C. Du [12] proposed silent position UW-ASN termed as UPS. This technique uses 4 anchor nodes to find location which is based on time difference of arrival. As anchor A is selected as master anchor which initiate the localization. It sends the beacon signal to B anchor and sensor nodes. B anchor reply the time of arrival of beacon signal and transmission time of its beacon signal. Then B, C and other anchor repeat the same process. Sensor nodes hear it and calculate the TDoA of beacons. It converts multiply the TDoA with speed of sound to calculate the range difference. Sensor nodes know the location of anchor nodes and do the self localization. This technique has low overhead, require no special hardware and provide silent positioning [13].

H. Tan, A.F. Gabor, Zahi Ang Eu, and W.K. G. Seah [14] proposed a wide coverage positioning termed as WPS. UPS technique is not able to find the locality of all sensor nodes. Sensor nodes which are near to anchor node need five anchors for localization which is done in [12]. WPS is same as UPS but it uses five anchors instead of four anchors. WPS is more accurately localize than UPS and also included the timeout.

TABLE I LOCALIZATION TECHNIQUES

Localization Technique	Anchor Property	Merits	Issues
AUV-Aided Localization	Propelled mobile anchor (AUV)	It Exploit the mobility of AUV to overcome the lack of GPS and sensor nodes can communicate in disconnected network	Battery of AUV drained very fast. There is high localization delay due to slow speed of AUV
Localization with directional beacons	Propelled mobile anchor (AUV)	Directional localization provides accurate localization.	AUV is restricted to float over the UASN which is not possible in practice
Dive 'N' Rise Localization	Non-Propelled mobile anchor.	It is less expensive and can localize 100% nodes with small errors	Deeper node get massage later than the node closer to the surface and it increase the localization delay
Large Scale Hierarchal localization	Underwater anchors, Surface buoys and reference node	It achieves high localization coverage with low error and low communication cost.	It has highest energy consumption and large overhead of exchanging beacons and messages.
Localization with Detachable Elevator Transceiver (DETL)	Underwater anchors, Surface buoys attached DETs and reference node	It decreases the cost of system and increase scalability.	It has large overhead.
Node discovery and network discovery localization technique	Anchor free	Local coordination system can make without need of anchor node.	Node discovery process take long time and it has high overhead and high energy consumption.
Silent position in Under water Sensor Network (UPS)	Four stationary anchors	UPS require no time synchronization and provide location privacy at under water sensor nodes.	It cannot localize the nodes which are outside area of four anchors.
Wide Coverage Positioning system(WPS)	Five stationary anchors	It claims high localizability space, high localization latency and low energy consumption as compared to UPS.	Its localization delay and communication cost is high as compared to UPS
Time Synchronization free localization in large scale under water(LSLS)	Stationary anchors	LSLS increase coverage of UPS by adding iterative phase.	LSLS has higher overhead and higher energy consumption than UPS.
Multi Anchor Node Collaborative Localization(MNCL)	Stationary anchors	MNCL increase the localization with low average error ration and low average energy consumption. It is more efficient than LSLS.	If ordinary nodes send more localization request messages, average energy consumption increases.

Wei Cheng, A. Thaeler, Xiuzhen Cheng, Fang Liu, X. Lu, and Zexin Lu [15] proposed a localization scheme in large scale under water termed as LSLS. This technique uses four anchor nodes as UPS but it adds iterative localization phase and complementary phase. Firstly it done localization as in UPS, then in iterative phase it select some nodes as reference nodes and help to localize other sensor nodes and repeat the localization process iteratively. Finally in complementary phase, un-localized nodes initiate localization and select different set of nodes as reference nodes and repeat localization.

Chenyu Zhang, Guangjie Han, Jinfang Jiang, Lei Shu, Guogao Liu, and Joel J.P.C Rodrigue [16] proposed multi anchor node collaborative localization (MNCL). MNCL divide the whole process into five sub-processes:-1) Ordinary node process. 2) Ordinary node localization process. 3) Iterative location estimation process. 4) Improved 3D Euclidean distance estimation process. 5) 3D DV-Hop distance estimation process based on two-hop anchor nodes. There are three types of nodes Surface buoys, anchor nodes and ordinary nodes. Monitoring area is divided into small cubes which are accomplished by surface buoys, each anchor node belong to one small cube region. Collaborative algorithm is applied to one cube region and to multiple cube regions. In third sub process temporary position of ordinary nodes is determined and in fourth sub process two hop anchor node help to localization with ordinary nodes.

By considering the above techniques, issues and merits of localization techniques are summarized in Table I.

IV. OPEN ISSUES AND FUTURE SCOPE

Various localization techniques have been reported in the previous section related to the UW-ASN. Considering the merits and opportunities available in the approach of localization of UW-ASN, open issues and future scope has been presented here:

Due to limited memory the recorded during the monitoring of aquatic environment is very limited. Any instrument failure fails the complete monitoring mission. Various instruments used in acoustic sensor network are very expensive. Acoustic sensor network need the location of some nodes should be known which practically very difficult task is.USN has very limited bandwidth. There is high bit error rate.

Instruments of USN usually fail due to corrosion. Use of cross layer approach is still an open issue. In future to overcome the various issue new techniques can be adopted or existing techniques can be improve in some manners. As cross layer approach can overcome the energy consuming problem of various nodes. Link quality should be considered to improve the accuracy. There is also need to consider the impact of localization techniques on location based protocols.

V. Conclusion

UW-ASN localization techniques play crucial role to find the location of sensor nodes and collect the monitoring data. Numerous localization techniques are proposed. This paper surveyed the UW-ASN localization techniques, merits and issues of those techniques. All these techniques have contributed to localize the sensor nodes but still there is a scope of holistic approach to solve the issues concerned. It is hoped that various issues highlighted here will prove to be helpful for the researchers working in this area.

References

[1] Ian F. Akyildiz, Dario Pompili, and Tommaso Melodia, " Underwater acoustic sensor networks: research challenges," *Ad Hoc Networks Journal (Elsevier)*, 2005.

[2] Melika Erol, Luiz F.M. Vieira, and Mario Gerla, "Localization with dive'n'rise (dnr) beacons for underwater acoustic sensor networks," *Second workshop on Underwater networks (WuWNet)*, Montreal, Quebec, Canada, 2007.

[3] Melika Erol, Luiz F.M. Vieira, and Mario Gerla, "Auv-aided localization for underwater sensor networks," *International Conference on Wireless Algorithms, Systems and Applications (WASA)*, Chicago, IL, US, 2007.

[4] Hanjiang Luo,Yiyang Zhao, Zhongwen Guo, Siyuan Liu, Pengpeng Chen, and Lionel M. Ni, "UDB: Using Directional Beacons for Localization in Underwater Sensor Networks," *14th IEEE International Conference on Parallel and Distributed Systems,* 2008.

[5] Kai-Ten Feng, "LMA: Location and mobility-aware medium access control protocols for vehicular ad-hoc networks using directional antennas," *Vehicular Technology,IEEE Transaction* , Vol. 56, Issue 6, Nov. 2007.

[6] Zhong Zhou, Jun-Hong Cui, and Shengli Zhou, "Localization for large-scale underwater sensor networks," *IFIP Networking*, Atlanta, Georgia, USA, pp. 108–119, May 2007.

[7] Melike Erol-Kantarci, Sema Oktug, Luiz Vieria, and Mario Gerla, "Performance evaluation of distributed localization techniques for mobile underwater acoustic densor networks," *Ad Hoc Netwoks*, Vol. 9,Issues 1, Jan 2011.

[8] Kai Chen, Yi Zhou, and Jianhua He, "A localization scheme for underwater wireless sensor networks," *International Journal of Advanced Science and Technology*, Vol. 4, March 2009.

[9] Yi Zhou, Bao-Jun Gu, Kai Chen, and Jian-bo Chen, and Hai-bing Guan, "An range-free localization scheme for large scale underwater wireless sensor networks," *Journal of Shanghai Jiaotong University*, Vol. 14,No. 5, pp. 562-568, October 2009.

[10] A.K. Othman, A.E. Adams, and C.C. Tsimenidis, "Node discovery protocol and localization for distributed underwater acoustic networks," *International Conference on internet, web applications and services (AICT-ICIW)*, Washington, DC, USA, Feb. 2006.

[11] A. K. Othman, "GPS-less localization protocol for underwater acoustic networks," *5th IFIP International Conference on Wireless and Optical Communications Networks*, pp. 1-6, 2008.

[12] Xizuhen Cheng, Haining Shu, Qilian Liang, and D. H.-C. Du, "Silent positioning in underwater acoustic sensor networks," *Vehicular Technology, IEEE Transaction*, Vol. 57, Issue. 3, pp. 1756-66, May 2008.

[13] Xizuhen. Cheng, Haining Shu, Qilian Liang, "A range-difference based selfpositioning scheme for underwater acoustic sensor networks," *International Conference on Wireless Algorithms, Systems and Applications (WASA)* ,Chicago, IL, US, pp. 38–43, 2007.

[14] H. Tan, A. F. Gabor, Zahi Ang Eu, and W. K. G. Seah, "A wide coverage positioning system (wps) for underwater localization," *IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, pp. 1–5,May 2010.

[15] Wei Cheng, A. Thaeler, Xiuzhen Cheng, Fang Liu, X. Lu, and Zexin Lu, "Time-synchronization free localization in large scale underwater acoustic sensor networks," 29th *IEEE conference on Distribute Computing Systems Workshops*, Montreal,QC, pp. 80-87, June 2009.

[16] Chenyu Zhang, Guangjie Han, Jinfang Jiang, Lei Shu, Guogao Liu, Joel J.P.C Rodrigue, "A Collaborative Localization algorithm for underwater acoustic sensor networks," *International conference on Computing, Management and Telecommunications (ComManTel)*, Da Nang, Vietnam, pp. 211 –216, April 2014.

Performance Analysis of AODV for Wormhole Attack Using Different Mobility Models

Gurmeet Kaur[1] and Amanpreet Kaur[2]
[1,2]*Centre for Computer Science and Technology, Central University of Punjab, Bathinda, India*
E-mail: [1]gurmeetpatil@gmail.com, [2]pandheraman@gmail.com

***Abstract*— Mobile Ad-Hoc Network (MANET) is a type of temporary wireless network, in which the nodes are mobile and have dynamic network topology. Communication among nodes in these networks is accomplished via different routing protocols. But these protocols have different security flaws and using these flaws, an attacker can launch many attacks. Wormhole attack is one of the serious attacks in the context of mobile ad-hoc networks that can disrupt any routing channel completely. In this work, an attempt has been made to analyze and compare the performance of on-demand reactive routing protocol: Ad hoc On Demand Distance Vector (AODV) with two approaches: AODV without attack and AODV under wormhole attack using two mobility models viz. Random Way Point Model and Reference Point Group Mobility Model. The performance metrics evaluated for the two examined approaches are Average Throughput, Packet Delivery Ratio, Average End to End Delay and Jitter. Along with this, an approach has also been used to analyze participated malicious nodes.**

Keywords: AODV, Mobility, RPGM, RWP, Tunnel

I. Introduction

Mobile Ad-Hoc Network (MANET) is emblematic and ubiquitous in nature, which is the extension of wireless networks. According to structural arrangement, wireless networks are classified into two main categories: fixed infrastructure wireless networks and infrastructure less wireless networks. Mobile Ad-Hoc Networks (MANETs) fall under the category of infrastructure less wireless networks [13] [3].

The original idea of MANET started out in the early 1970s and during this period of time, MANET was called “packet radio” network sponsored by DARPA. The whole life cycle of ad hoc networks could be categorized into three generations and present ad hoc networking systems are considered the third generation, which was started out in 1990s [23].

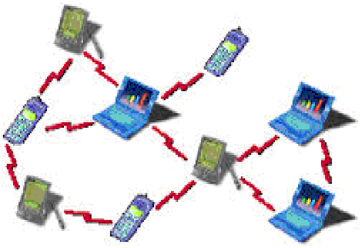


Fig. 1 Infrastructure Less Wireless Network (Mobile Ad-hoc Network) [1]

There are various MANET routing protocols as no single routing protocol works well in all environments [15]. The reason is that the traditional routing protocols (which have already written for the wired network) do not perform well in MANETs. Hence there was a need to design new protocols for mobile ad hoc networks [24].

Depending upon the many ways by which computers can communicate, the routing protocols in mobile ad-hoc network can be divided into three categories [18]: Demand Oriented, Table Oriented and Hybrid Routing Protocols. But in this work, a demand oriented or reactive routing protocol is used for analysis: AODV (Ad Hoc On-Demand Distance Vector).

The first version of AODV was published in November 2001 by Working Group for routing of the IETF community. It uses sequence numbers to solve the count-to-infinity and loop creation problem [16]. It includes two main steps for the proper working namely Route Discovery and Route Maintenance with the help of four types of control messages: RREQ, RREP, RERR and HELLO [12].

The remaining paper is organized into various sections as follows: section II gives the brief introduction to mobility models, which are used in simulation process. Section III illustrates the scenario of wormhole attack in AODV protocol and its types. In section IV, the simulation environment and methodology is explained. Section V lists the various results. Section VI provides the conclusion followed by references.

II. Mobility Models

Nowadays, for the simulation of realistic movements that are produced by users of a mobile or wireless network, different mobility models are used [20]. There are two main categories of mobility models namely purely synthetic models and trace-based mobility models [22]. But amongst them, purely synthetic models are commonly used in research. In this research work random waypoint and reference point group mobility models are used.

A. Random Way Point Model (RWP)

It was first proposed by Johnson and Maltz. It is elementary synthetic model, which is used to evaluate the MANET routing protocols [6] [7]. In this model, at every instant, a node randomly chooses a destination anywhere in the specified network field and moves towards it with a velocity chosen randomly from a uniform distribution between {0, V_max}, where V_max is the maximum allowable velocity for every mobile node. To overcome sudden stop and start, ‘pause time’ parameter is used by nodes. For this duration, the node stops after reaching the destination. Fig. 2 illustrates an example of a topography showing the movement of nodes for RWP [8] [17].

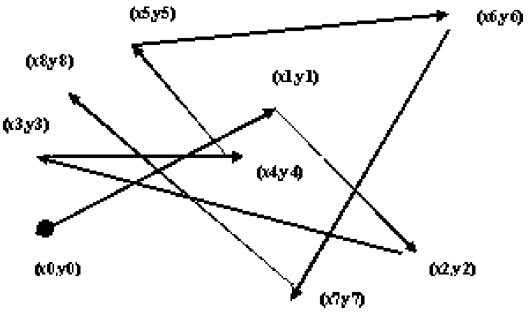


Fig. 2 Node Movement in Random Way Point Model [7]

B. *Reference Point Group Mobility Model (RPGM)*

In contrast to RWP, there are some spatial dependencies in RPGM. To simulate the group movement behavior in the real world such as communication in military battlefield and disaster areas, reference point group mobility model was proposed [5].

In RPGM, nodes are divided into groups or clusters. Each group has a logical center called group leader that defines the whole group’s motion behavior and leader’s mobility follows random waypoint. Initially each member of the group is uniformly distributed in the neighborhood of the group leader. Then, at each instant, every node has speed and direction that is derived by randomly deviating from that of the group leader [21].

Figure 3 shows an example topography illustrating the movement of nodes for Reference Point Group Mobility Model.

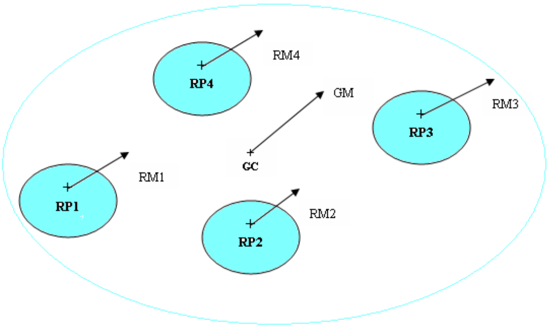


Fig. 3 Reference Point Group Mobility Model [21]

where, RP: Random Point
RM: Random Motion
GC: Group Center
GM: Group Movement

III. Wormhole Attack & its Types

Security in MANET plays a vital role for basic network functions. Availability, Authorization, Confidentiality, Integrity and Non-repudiation are some basic requirements that effective security architecture must ensure in order to combat passive and active attacks [15] [2] [14].

A. *Wormhole Attack in AODV Protocol*

According to [9] [19] wormhole attack is an active attack. Wormhole attacker affects the original functionality of MANET routing protocols such as AODV, DSR and OLSR etc, but this research work emphasizes on wormhole attack in AODV routing protocol. A simplified view of wormhole attack is shown in Fig. 4.

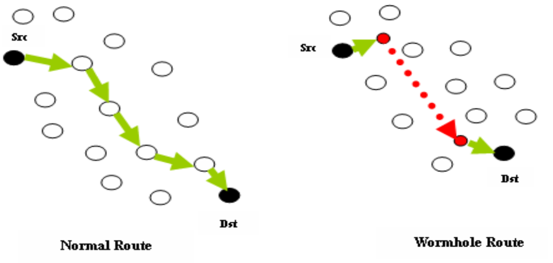


Fig. 4 Scenario of Wormhole Attack [4]

Suppose a source wants to talk with destination. And this communication is possible through shortest path provided by AODV protocol (called normal route). But if two malicious nodes are kept at two different locations in the network and a malicious node accepts the traffic at one location, tunnels them through wormhole link to another malicious node, then replays packets into the network at that location, then this is called wormhole route [10].

Hence, the functioning of AODV protocol is completely disrupted by this attack. It affects various QoS parameters such as delay, jitter, throughput, packet delivery ratio and energy consumption etc [9] [4]. The various metrics such as strength, packet delivery ratio, path length, attraction and robustness etc can also be used to detect wormhole attack in the network [25].

B. *Types of Wormhole Attack*

According to [6], there are various types of wormhole attack as follows:

1. *All Pass:* Here malicious nodes can pass all the packets regardless of their size.
2. All Drop: Here malicious nodes can drop all the received packets in the network.
3. *Threshold:* Sometimes, there is a constraint as a threshold value in network and malicious node can drop all the packets having size greater than or equal to the threshold value.

4. *Replay:* Here, one malicious node can replay the packets after tunnelling in the network.
5. *Tunnelling:* Wormhole attack is also called tunnelling attack. So here, a malicious node tunnels the packets from one location to another location in the network via wormhole link.
6. *Propagation Delay:* The propagation delay in the network is increased as more time is taken by malicious nodes to send packets from source to destination.

IV. SIMULATION SETUP & METHODOLOGY

To construct a real distributed testing environment, the cost and complexity is very high. So simulation is widely used in network research. Simulation is the manipulation of the model of a system that is used to observe the behavior of a particular system in a setup similar to real-life [11]. NS2 simulator is used in this research work and it is the most widely used simulator in academia. This study was performed on Intel Core i7 computer system using Ubuntu Linux 12.04 Operating System.

A. *Simulation Methodology*

This work has been divided into following steps:

Step 1: Simulation of the demand-oriented routing protocol AODV under two synthetic mobility models: RWP (Random Way Point) and RPGM (Reference Point Group Model).

To simulate AODV under random waypoint mobility model, a number of nodes (from 10-50) are uniformly distributed in an area size of 1186*584 sq. m. having CBR traffic type. And to simulate AODV under reference point group mobility model, five configurations with different number of nodes have been configured as follows:

- *Configuration I:* When network size is small i.e. network is having 10 nodes only, then 1 group with 10 nodes is configured.
- *Configuration II:* For 20 nodes, 2 groups are configured with 10 nodes each.
- *Configuration III:* For 30 nodes, 3 groups are configured with 10 nodes each.
- *Configuration IV:* For 40 nodes, 4 groups are configured with 10 nodes each.
- *Configuration V:* For 50 nodes, 5 groups are configured with 10 nodes each.

The movement scenarios of nodes for both mobility models are generated through bonnmotion tool.

Step 2: Simulation of AODV under wormhole attack using two mobility models: RWP (Random Way Point) and RPGM (Reference Point Group Model).

To simulate wormhole attack, malicious nodes are kept at different locations in the already created topology for both mobility models and the required coding is done to create wormhole tunnel with the help of other nodes in the network, which bypass normal route. In this scenario, minimum number of malicious nodes is 1, but tunnel length increases as network size increases.

Step 3: Graphical analysis and performance comparison of normal AODV and AODV under attack environment using RWP and RPGM.

Using AWK scripts, various performance metrics such as PDR, average throughput, jitter and average end to end delay have been analyzed graphically and comparison is done between AODV without attack and AODV under attack by varying number of nodes.

Step 4: Analysis of the malicious nodes which are participating to make wormhole peer list in the network.

To analyze malicious nodes, an implementation has been done at NS2 link layer. Required coding has been done in ll.cc and ll.h files at link level. Firstly, in ll.cc and ll.h files, parameters such as size of wormhole peer list (tunnel) and properties of nodes are defined and then in Tcl file, the definition of nodes is configured. During this analysis, the tunnel length varies from 1 to 5 nodes.

The simulation parameters for all above steps are shown in Table 1:

TABLE I SIMULATION PARAMETER

Parameters	Value
Simulator	NS-2 Version 2.35
Number of Nodes	10, 20, 30, 40, 50
Topography Dimension (m*m)	1186*584
Simulation Time	90 seconds
Traffic Type	CBR
Signal Propagation Model	Two Ray Ground Model
MAC Type	802.11 MAC Layer
Data Rate	2.0 Mb
Mobility Models	Random Waypoint, Reference Point Group
Routing Protocol	AODV
Interface Queue	Drop Tail/Priority Queue
Channel	Wireless Channel
Link Layer Type	LL
Antenna Type	Omni direction
Minimum Number of Malicious Nodes	1
Tunnel Length	1-5 nodes
Probability of Group Change	0.01
Maximum Distance between Groups	1.0
Average Number of Nodes in a Group	10
Min Speed and Max Speed of Nodes	0.5 and 1.5 m/s
Performance Metrics	PDR, Average Throughput, Average End to End Delay and Jitter
Examined Approaches	without attack and under attack

V. Results & Discussion

A. *Performance Analysis of AODV Protocol under RWP and RPGM*

AODV Protocol is simulated by varying number of nodes using CBR traffic and two mobility models.

1) Average throughput

The throughput tends to fluctuate with the increase in network size under random waypoint model. Reference point group mobility model offers higher throughput than the random waypoint.

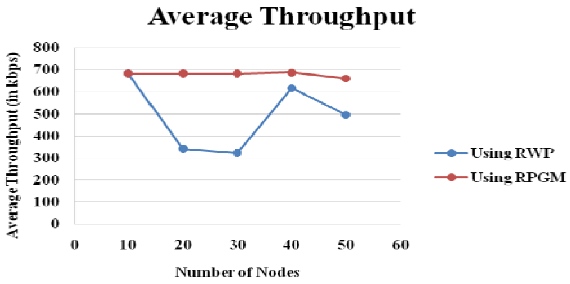


Fig. 5 Average Throughput of AODV under Different Mobility Models

2) Average end to end delay

Random waypoint model exhibits lesser delay than the reference point group mobility model. Due to configuration of various groups in reference point group model, delay is high in case of RPGM, but value of delay decreases as number of nodes or groups increases in RPGM.

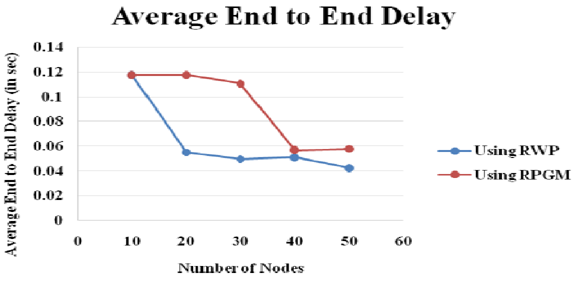


Fig. 6 Average End to End Delay of AODV under Different Mobility Models

3) Packet delivery ratio

The packet delivery ratio decreases with increase in network size. And the value of PDR under RPGM is more as compared to RWP. Initially, packet delivery ratio remains constant for network size of 10 and 20 in RPGM, but then decreases gradually.

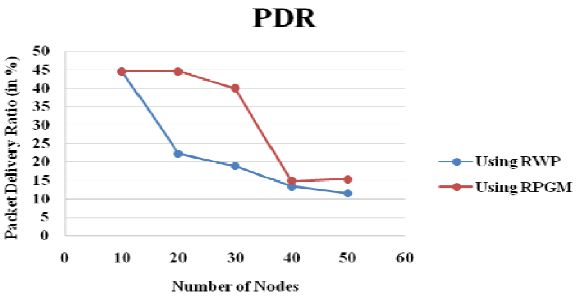


Fig. 7 Packet Delivery Ratio of AODV under Different Mobility Models

4) Jitter

There is a fluctuation in jitter graph for RWP. But for RPGM, the values are nearly same as number of nodes increases up to 40 nodes, but after that it decreases. Overall jitter is high in case of RPGM.

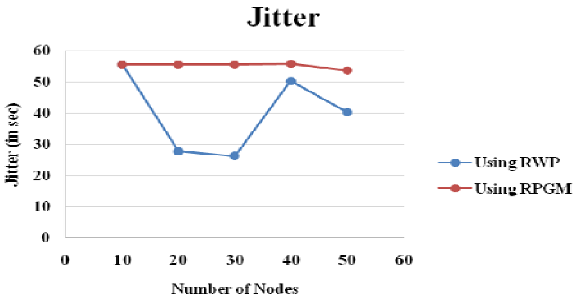


Fig. 8 Jitter of AODV under Different Mobility Models

B. *Performance Analysis of AODV under Attack Environment using RWP and RPGM*

1) Average throughput

The value of throughput decreases as network size and tunnel length increases in case of RPGM. But in RWP, throughput varies.

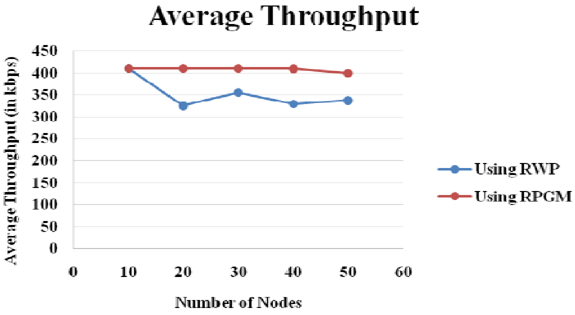


Fig. 9 Average Throughput of AODV under Attack using Different Mobility Models

2) *Average end to end delay*

There is a variation in delay in both scenarios. But delay increases as network size and tunnel length increases. Overall RPGM exhibits more delay due to increase in network size, tunnel length and number of groups.

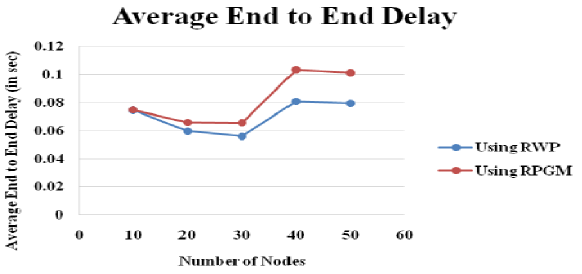


Fig. 10 Average End to End Delay of AODV under Attack using Different Mobility Models

3) *Packet Delivery Ratio*

Initially, the value of PDR decreases up to 30 nodes and then suddenly increases for 40 nodes and then again decreases. The sudden increase is due to the tunnelling and replaying nature of attack. More tunnel length and replay, more packets will be delivered.

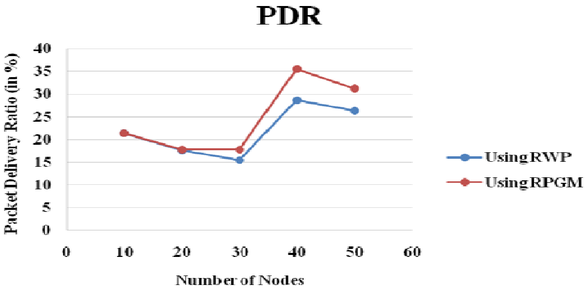


Fig. 11 Packet Delivery Ratio of AODV under Attack using Different Mobility Models

4) *Jitter*

Jitter is more in case of reference point group mobility model and it remains almost same up to 40 nodes and then decreases. But in case of random waypoint model, initially jitter is high and then variation starts.

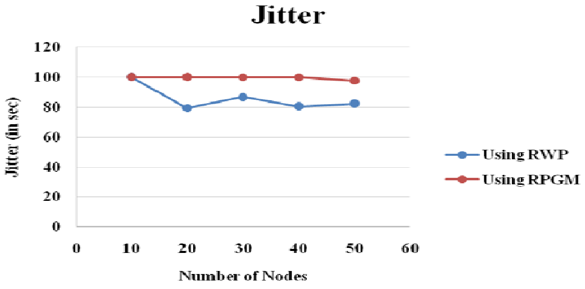
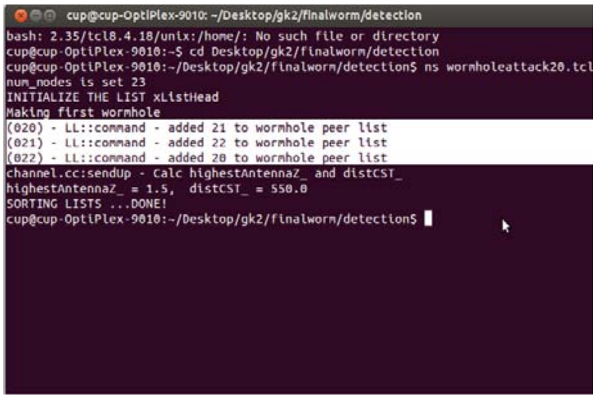


Fig. 12 Jitter of AODV under Attack using Different Mobility Models

C. *Analysis of Malicious Nodes in Wormhole Peer List*

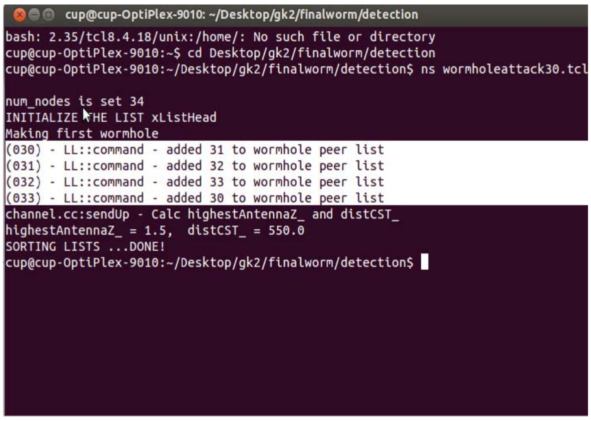
Fig. 13 shows that malicious nodes 20, 21 and 22 are participating to make tunnel (having tunnel length 3 nodes) and disrupt the normal path of AODV protocol.



```
cup@cup-OptiPlex-9010: ~/Desktop/gk2/finalworm/detection
bash: 2.35/tcl8.4.18/unix:/home/: No such file or directory
cup@cup-OptiPlex-9010:~$ cd Desktop/gk2/finalworm/detection
cup@cup-OptiPlex-9010:~/Desktop/gk2/finalworm/detection$ ns wormholeattack20.tcl
num_nodes is set 23
INITIALIZE THE LIST xListHead
Making first wormhole
(020) - LL::command - added 21 to wormhole peer list
(021) - LL::command - added 22 to wormhole peer list
(022) - LL::command - added 20 to wormhole peer list
channel.cc:sendUp - Calc highestAntennaZ_ and distCST_
highestAntennaZ_ = 1.5,  distCST_ = 550.0
SORTING LISTS ...DONE!
cup@cup-OptiPlex-9010:~/Desktop/gk2/finalworm/detection$
```

Fig. 13 Analysis of Three Malicious Nodes

Similarly, Fig. 14 shows that malicious nodes 30, 31, 32 and 33 are participating to make tunnel (having tunnel length 4 nodes) and disrupt the normal path of AODV protocol.



```
cup@cup-OptiPlex-9010: ~/Desktop/gk2/finalworm/detection
bash: 2.35/tcl8.4.18/unix:/home/: No such file or directory
cup@cup-OptiPlex-9010:~$ cd Desktop/gk2/finalworm/detection
cup@cup-OptiPlex-9010:~/Desktop/gk2/finalworm/detection$ ns wormholeattack30.tcl

num_nodes is set 34
INITIALIZE THE LIST xListHead
Making first wormhole
(030) - LL::command - added 31 to wormhole peer list
(031) - LL::command - added 32 to wormhole peer list
(032) - LL::command - added 33 to wormhole peer list
(033) - LL::command - added 30 to wormhole peer list
channel.cc:sendUp - Calc highestAntennaZ_ and distCST_
highestAntennaZ_ = 1.5,  distCST_ = 550.0
SORTING LISTS ...DONE!
cup@cup-OptiPlex-9010:~/Desktop/gk2/finalworm/detection$
```

Fig. 14 Analysis of Four Malicious Nodes

VI. Conclusion

The performance analysis of AODV without attack and under attack has been carried out in a comprehensive manner using random waypoint and reference point group mobility models.

Firstly, AODV without attack is analyzed under random waypoint and reference point group models. Results show that AODV performs well for throughput, PDR and packet drop rate under RPGM and for delay and jitter under RWP. Secondly, AODV under wormhole attack is analyzed using two mobility models namely random waypoint and reference point group models. Analysis shows that AODV under attack gives high value for throughput, PDR, delay and jitter in RPGM and low for packet drop rate in RWP.

Along with above, one more step has been taken to analyze the malicious nodes which are making tunnel to perform attack.

REFERENCES

[1] ACoRN, “Ad Hoc Networks”, *ARC Communications Research Network*, 2010, Available at: http://www.acorn.net.au/telecoms/adhocnetworks/adhocnetworks.html.

[2] B. Kannhavong, H. Nakayama, Y. Nemoto, N. Kato, and A. Jamalipour, “A Survey of Routing Attacks in Mobile Ad Hoc Networks”, *IEEE Wireless Communications*, vol. 14, no. 5, pp. 85-91, Oct. 2007.

[3] C.E. Perkins, “Ad Hoc Networking with AODV”, 2013, Available at: http:// www.psg.com/~charliep/txt/Daedeok2002/AODV-Daedeok.pdf.

[4] C. P. Vandana, and A. F. S. Devaraj, “Evaluation of Impact of Wormhole Attack on AODV”, *International Journal of Advanced Networking and Applications*, vol. 4, no. 4, pp. 1652-1656, 2013.

[5] D. Kumar, A. Srivastava, and S. C. Gupta, “Routing in Ad Hoc Networks under Reference Point Group Mobility”, *European Modelling Sysposium, IEEE Computer Society*, pp. 595-598, 2013.

[6] F. A. Jenefer, and D. Vydeki, “Performance Analysis of Mobile Ad Hoc Network in the Presence of Wormhole Attack”, *International Journal of Advanced Computer Engineering and Communication Technology*, vol. 1, no. 1, pp. 13-18, 2013.

[7] F. Bai, and A. Helmy, “Chapter 1: A Survey of Mobility Models”, pp. 1-30, Available at: www.cise.ufl.edu/~helmy/papers/Survey-Mobility-Chapter-1.pdf.

[8] F. Bai, N. Sadagopan, and A. Helmy, “User Manual for IMPORTANT Mobility Tool Genarators in NS-2 Simulator”, pp. 1-12, Feb. 2004.

[9] G. K. Singh, A. Kaur, and A. L. Sangal, “Performance Analysis of DSR, AODV Routing Protocols based on Wormhole Attack in Mobile Ad-hoc Network”, In *Proceedings of the 5th IEEE International Conference on Advanced Computing & Communication Technologies*, pp. 31-36, 2011.

[10] G. Kaur, and A. Kaur, “A Comprehensive Review on Performance of AODV Protocol for Wormhole Attack”, *International Journal of Research in Engineering and Technology*, vol. 3, no. 5, pp. 531-537, May 2014.

[11] J. Singh, K. Kumar, M. Sachdeva, and N. Sidhu, “DDoS Attack’s Simulation using Legitimate and Attack Real Data Sets”, *International Journal of Scientific & Engineering Research*, vol. 3, pp. 1-5, June 2012.

[12] N. Gandhewar, and R. Patel, “Detection and Prevention of Sinkhole Attack on AODV Protocol in Mobile Ad Hoc Network”, In *Proceedings of the Fourth International Conference on Computational Intelligence and Communication Networks, IEEE Computer Society*, pp. 714-718, 2012.

[13] N. Khemariya, and A. Khuntetha, “An Efficient Algorithm for Detection of Blackhole Attack in AODV based MANETs”, *International Journal of Computer Applications*, vol. 66, pp. 18-24, March 2013.

[14] P. G. Argyroudis, and D. O’Mahony, “Secure Routing for Mobile Ad Hoc Networks”, IEEE Communications Surveys & Tutorials, Third Quarter, vol. 7, no. 3, pp. 2-21, 2005.

[15] R. Agrawal, R. Tripathi, and S. Tiwari, “Performance Comparison of AODV and DYMO MANET Protocols under Wormhole Attack Environment”, *International Journal of Computer Applications*, vol. 44, no. 9, pp. 9-16, April 2012.

[16] R. Baumann, “AODV: Ad hoc On Demand Distance Vector Routing Protocol”, pp. 1-19, April 2002, Available at: http://www.rainer-baumann.ch/public/qec.pdf.

[17] R. Mohan, C. Rajan, and N. Shanthi, “A Stable Mobility Model Evaluation Strategy for MANET Routing Protocols”, *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 2, no. 12, pp. 58-65, Dec. 2012.

[18] S. Gandhi, N. Chaubey, N. Tada, and S. Trivedi, “Scenario-based Performance Comparison of Reactive, Proactive and Hybrid Protocols in MANET”, In *Proceedings of the IEEE International Conference on Computer Communication and Infomatics*, pp. 1-5. 2012.

[19] S. Gupta, S. Kar, and S. Dharmaraja, “WHOP: Wormhole Attack Detection Protocol using Hound Packet”, In *Proceedings of the* IEEE International Conference on Innovation Technology, pp. 226-231, 2011.

[20] S. Kumar, D. Singh, and M. Chawla, “Performance Comparison of Routing Protocols in MANET Varying Network Size”, *International Journal of Smart Sensors and Ad Hoc Networks*, vol. 1, no. 2, pp. 51-54, 2011.

[21] S. Kumar, S. C. Sharma, and B. Suman, “Impact of Mobility Models with Different Scalability of Networks on MANET Routing Protocols”, *International Journal of Scientific & Engineering Research*, vol. 2, no. 7, pp. 1-5, July 2011.

[22] T. Issariyakul, and E. Hossain, “Introduction to Network Simulator”, *Computer Networks*, 2011, Available at: books.google.co.in/books?isbn=1461414067.

[23] V. C. Patil, “Chapter-3: Overview of Mobile Ad Hoc Networks”, pp. 19-36, 2012, Available at: http://www.shodhganga.inflibnet.ac.in/bitstream/10603/4106/.../11_chapter%203.pdf.

[24] V. K. Upadhyay, and R. Shukla, “An Assessment of Worm Hole Attack over Mobile Ad-Hoc Network as Serious Threats”, *International Journal of Advanced Networking and Applications*, vol. 5, pp. 1858-1866, 2013.

[25] V. Mahajan, M. Natu, and A. Sethi, “Analysis of Wormhole Intrusion Attacks in MANETs”, In *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pp. 1-7. 2008.

Challenges to WiMAX QoS Scheduling and its Mitigation Schemes – A Review

Harsukhpreet Singh[1], Vaibhav Pandey[2] and Anita Mishra[3]
[1]CTIT, Jalandhar, PTU Jalandhar, Punjab, India
[2,3]PIT, PTU, Jalandhar, Punjab, India
E-mail: [1]harsukhpreet@gmail.com, [2]pandeyvaibhav51@gmail.com,
[3]anitamishramtech@gmail.com

***Abstract*—For realization of future high performance integrated networks, broadband distribution and access networks and to meet the increasing demand of multimedia services with a guaranteed quality of service, WiMAX becomes out as the most promising technology with flexibility and mobility of wireless networks. WiMAX has great impact in field of 4th generation networks with introduction of VoIP applications giving advantage to utilize existing infrastructure in the form of internet connection having several QoS scheduling algorithms such as UGS, Best Effort, rtPS, nrtPS and ertPS. In this paper, we review the different challenges that limit the QoS capabilities of VoIP-WiMAX communication networks and different mitigation techniques to combat with these challenges to realize high performance WiMAX using scheduling algorithms.**

Keywords: WiMAX, VoIP, QoS, QoS Scheduling, OFDM

I. INTRODUCTION

The increasing demand for larger capacity and higher transmission speeds to accommodate for data-intensive multimedia in conjunction with real-time applications, the wireless networks have experienced an explosive growth in last few years [1–2]. WiMAX stand for Worldwide Inter-portability for Microwave Access can be a communication technology for easily delivering high-speed data rates to large geographical areas using orthogonal frequency division multiplexing (OFDM) from Base Station (BS) to Subscriber Station (SS) which mitigates noise, multipath and interference effects, what exactly are primary challenges of wireless communication [3–4]. The WiMAX network is a combination of subscriber station (SS) and base station (BS). Here the packets are transferred from source node to destination node after following various scheduling, modulation technique and routing technique. In compliance with IEEE 802.16 the maximum range of WiMAX network is 50 km from the BS, where the responsibility of the base station for providing the air interface to the master station with supplementary functions that may be part of the BS are micro mobility management functions, radio resource management, handoff triggering, traffic classification, tunnel establishment, QoS policy enforcement, key management, DHCP proxy, multicast group management and session management. The receiver and antenna could be a small box or Personal Computer Memory Card International Association (PCMCIA) card or a laptop as shown in Fig. 1 [5–6].

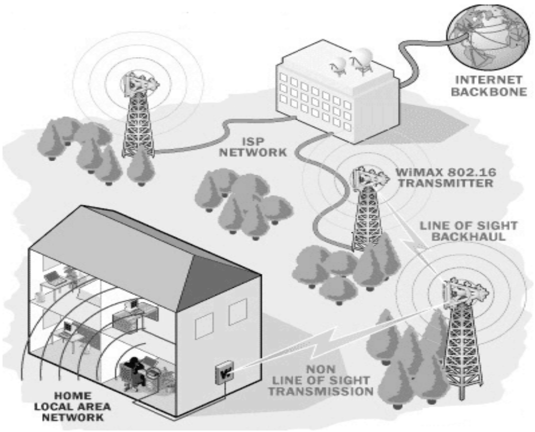


Fig. 1 WiMAX Network Architecture

With the introduction of 4G technology in WiMAX offering a metropolitan area network service that can use one or more base station and each base station provide the service to the users up to 50 km radius for spreading broadband wireless data over spacious geographic area. WiMAX offers high-speed, flexible, inexpensive and last mile services with performance similar to those of wire line infrastructures T1, DSL, cable modem based connections, optical fiber or copperware with a number of QoS needs. WiMAX provides wide area coverage and quality of service capabilities for applications ranging from real-time delay-sensitive Voice-over-Internet Protocol (VoIP) to real-time streaming video and non-real-time downloads, ensuring that subscribers obtain the performance they expect for all types of communications with a bandwidth support of upto 10 MHz [7–8]. Based on **IEEE 802.16 standard,** WiMAX provides upto 30 miles broadband access to mobile users having a telecommunication etiquette offering full access to mobile internet across cities and countries with a wide range of devices. WiMAX technology is offering very high speed **broadband access** to mobile internet. Generally 10MHz with the TDD scheme provides 3:1 up and down link ratio. The architecture of **WiMAX technology** based on **MAC layer** which is a connection oriented layer. Through MAC layer a use can perform a variety of functions such as various type of application including multimedia and voice can be used. It also supports best efforts for data traffic as bit, real time, traffic flaws etc. The aim to design WiMAX technology is to facilitate large number of user with a variety of connection per terminal [9–10].

This paper is divided into three sections. Some introductory application and characteristic areas of WiMAX technology are presented in Section 1followed by Section 2 that presents the previous research work on different QoS scheduling schemes of WiMAX technology, challenges and their mitigation schemes and finally, the conclusion is reported in Section 3.

II. Challenges and Mitigation Methods in WiMAX QoS Scheduling

With the rollout of the third generation cellular networks, the aim is already set towards the next generation. Future generation networks will be characterized by variable and high data rates, QoS services and seamless mobility both within a network and among different networks. The rapid growth of wireless communications is mainly attributed to their ease of installation in comparison to fixed networks [11–13]. QoS more barely refers to meeting certain necessity typically, packet delivery ratio, packet error rate, jitter, SNR (Upstream/Downstream) and delay associated with a given application and various demonstrations has been provided [14–16]. WiMAX networks must support a diversity of applications, such as video, voice, multimedia and data and each of these has different traffic patterns and QoS necessity and allows a better QoS handling [17]. QoS in WiMAX broadly divided in two parts, User-Centric in which the collective effect and services information performances work out how much user satisfaction is within the service and Network-Centric QoS and Network-Centric QoS comprises the systems that provide network managers and chance to manage this combination of bandwidth, delay, variances in delay (jitter) and packet reduction in the network to capable of deliver a network service [18]. In MANET, due to frequent variations in mobility of nodes in terms of speed, direction and rate, the structure of the network varies dynamically and unpredictably over time and causes route failures, which effects its QoS [19–20]. Whereas, in WiMAX the QoS is granted on the basis of type of application and service under consideration. For example, an user sending an email needs no real-time data stream like another user having a Voice over IP (VoIP) application. To provide the service parameters respectively, the traffic management is necessary. There are four main service classes named as UGS, rtPS, nrtPS, BE but there is a fifth type QoS service class which is added in 802.16e standard, named as extended real-time Polling Service (ertPS). Within all these scheduling resources are allocated to manage and satisfy the QoS of higher priority services [21–22]. Table 1 broadly classifies various service classes defined in WiMAX and its applications.

Table I QoS Classes in WiMAX [14]

Service Classes	Description	Applications
Unsolicited Grant service (UGS)	For constant Bit rate and delay dependent applications	VoIP
Real Time Polling Service (rtPS)	For variable rate and delay dependent applications	Streaming audio, video
Extended Real time Service (ertPS)	For variable rate and delay dependent applications	VOIP and Silence Suppression
Non real time polling service (nrtPS)	Variable and non real time applications	FTP
Best Effort (BE)	Best effor	Email, Web Traffic

III. Impact of QoS Scheduling and its Mitigation

IEEE 802.16 has five QoS classes which Unsolicited Grant Service (UGS) is design to support real time service flow which generates the fixed size data packet periodically. In this algorithms BS assign fixed size grant to the subscriber station. The grants assign are basically of two type i.e. grant size and grant period. When voice session is initialized then these values are conciliated. These grants are sufficient for sending data packets. UGC service minimize the MAC overhead and uplink access delay which are caused when SS make request to the BS for bandwidth request to send the voice data packets, but UGS assign fixed size grant for sending voice data packets but voice user do not always have voice data packet to send because they have period of silence and it cause a waste of uplink resources. Real time polling service (rtPS) is designed to support real time services, which generally generates variable size data packets periodically. BS assign uplink resources to the SS when voice session is initialized then these values are conciliated. In this algorithm the SS request the BS for bandwidth of suitable size grant so that the rtPS can transport data more efficiently as compare to UGS algorithms. Because the SS always made a request for Bandwidth to the BS which in turn can cause more MAC overhead and uplink access delay as compare to UGS algorithms [23–24]. Extended Real Time Polling Services (ertPS) algorithm is proposed in order to remove the shortcomings of both UGS and rtPS algorithms. The UGS approach allows BS to assign fixed size grants to voice users which leads to wastage of uplink resources during the silent period when voice users do not have any data to send. Meanwhile rtPS although meant for variable size data packets consumes much of the time in polling process and also account for MAC overhead. So ertPS is designed for applications with variable size data packets because it does not have much MAC overhead and access delay. Thus the approach will be suitable for real-time applications like VOIP [25–26].

In ertPS algorithm whenever voice users have to increase their bandwidth users inform their BS by sending Bandwidth Request Header while setting BR (Bandwidth Request) bits to 1 in order to distinguish them from general BR bits. In this case, BS assigns uplink resources according to the requested size until user requests for another change in bandwidth. Whereas if users want to decrease the bandwidth users inform their BS by sending Grant Management sub-header while setting PBR (Piggyback Request) bits to 1 in order to distinguish them from general PBR bits. Again, BS assigns uplink resources according to the requested size until user requests for another change in bandwidth [27–28]. The ertPS only follows the request made by users periodically to assign the uplink resources, thus the rate drops to half the original rate and again it remains the same till voice user requested yet another decrement in the requested bandwidth and so on. Thus ertPS proves to be a better approach than UGS and RTPS for dealing with data transfer in real-time services more efficiently and it does not include any overheads also [29–30]. The non real time polling service (nrtPS) approach allows BS to assign variable size grants to the voice users on regular basis. This service basically supports those applications which require high throughput like FTP (File transfer Protocol) but can bear delay. In this algorithm the BS station assign uplink resources to the SS station same as rtPS algorithms but the resources are provided at the longer intervals. This will ensure that during the network congestions the SS station receives the request opportunities [31–34]. Best Effort (BE) service approach is used to support with the purpose of data stream where no minimum service level is required and it is provide efficient service to best effort traffic. These flows served by contention slots [35].

IV. Conclusion

A comprehensive review of research in the area of WiMAX QoS scheduling and challenges that exist within WiMAX performance and strategies to mitigate these various impairments to enhance the overall link performance is presented in this work. The main focus of this work is to put attention towards the realization of future high performance WiMAX networks, integrated with mobile broadband distribution and access networks by reviewing the past few year efforts in the area of impairments associated with WiMAX links and its mitigation techniques.

References

[1] Vishal Sharma, Harsukhpreet Singh, Jagjit Malhotra, "Performance Analysis of IEEE 802.11e (EDCF) and IEEE 802.11(DCF) WLAN Incorporating Different Physical Layer Standards", Institution of Engineers: Series B, Springer, December 2012, Volume 93, Issue 4, pp 247–253

[2] Vishal Sharma, Harsukhpreet Singh, Jagjit Malhotra, "Performance Evaluation of MAC- and PHY-Protocols in IEEE 802.11 WLAN", In proceeding of HPAGC (2011), Springer, CCIS 169, pp. 497–503, 2011.

[3] Andrews; J.G; A. Ghosh and R. Muhamed, 2007. Fundamentals of WiMAX: Understanding Broadband Wireless Networking. Prentice Hall, Englewood Cliffs, NJ.

[4] Chakraborty, M. and D. Bhattacharyya, 2010, "Overview of End-to-end WiMAX Network Architecture', WiMAX Security and Quality of Service, Head, Department of Information Technology, Institute of Engineering and Management, Salt Lake, Kolkata, India.

[5] Nee, R.V. and R. Prasad, 2009. OFDM for Wireless Multimedia Communications. Artech House Publishers, University of Michigan U.S.A., pp: 124–130.

[6] R. K. Jha, A. V. Wankhede and U. D. Dalal, "A Survey of Mobile WiMAX IEEE 802.16 m, " *International Journal of Computer Science and Information Security*, Vol. 8, No. 1, 2010, pp. 125–131

[7] Grondalen, O.; Gronsund, P.; Breivik, T.; Engelstad, P., "Fixed WiMAX Field Trial Measurements and Analyses, " *Mobile and Wireless Communications Summit, 2007. 16th IST*, vol. no. 1, pp.1–5, July 2007 doi: 10.1109/ISTMWC.2007.4299213.

[8] Grewal, V. and A.K. Sharma, 2010. On performance evaluation of different QoS mechanisms and AMC scheme for an IEEE 802.16 based WiMAX network. *Int. J. Comp. Appl*., 6(7).

[9] IEEE Std. 802.16, 2004 . IEEE Standard for local and metropolitan area networks Part16: Air Interface for Fixed Broadband Wireless Access Systems. ISBN: 0-7381-4070-8.

[10] Iwan Adhicandra, "Measuring Data and VoIP Traffic in WiMAX Networks", *journal of telecommunications*, volume 2, issue 1, april 2010.

[11] P. Delannoy, H.D. Nguyen, M. Marot, N. Agoulmine, M. Becker, "WiMax quality-of-service estimations and measurement." LRSM/IBISC Laboratory, University of Evry Val d'Essonne, 1 boulevard Francois Mitterrand, 91000 Evry, France, CNRS SAMOVAR Laboratory, UMR 5157, Telecom Paris-Sud, France.

[12] Vishal Sharma, Harsukhpreet Singh, Shashi Kant, "AODV based energy efficient IEEE 802.16g VANET", Fifth International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom 2013), The Institute of Engineering & Technology (IET), January 2013, pp: 35–43, DOI: 10.1049/cp.2013.2221; ISBN: 978-1-84919-842-4

[13] Vishal Sharma, Mandeep Kaur, Vijay Banga, "Performance Evaluation of Reactive Routing Protocols in MANET Networks using GSM based Voice Traffic Applications", Optik, Elsevier, Volume 124, Issue 15, August 2013, Pages 2013–2016

[14] Vishal Sharma, Navneet Kaur, "Location Based QoS Estimation of OFDM-WIMAX Network", DOI WC122012005, CiiT-International Journal of Wireless Networking, Dec 2012

[15] Vishal Sharma, Navneet Kaur, "Performance Estimation of OFDM-WIMAX Network", IJCNWMC, Trans steller, volume 3, Issue 1, pp: 39–46, March 2013

[16] Vishal Sharma, Navneet Kaur, "Protocol based QoS Estimation of OFDM-WIMAX Network", International Conference on Recent Trends in Communication and Computer Networks (COMNET'13), Elsevier, Nov 08-09, 2013; Hyderabad, India, ComNet2013-550R

[17] Safa, M. and M. Watfa, "Quality of Service in Wireless Local and Metropolitan Area Networks, " Auerbach Publications, CRC Press, Taylor and Francis Publishers, 2009.

[18] Upena D.Dalal and Y.P.Kosta, " WiMAX New Development, " 2009.

[19] Vishal Sharma, Shashi Kant, Harsukhpreet Singh, "Implementation and Analysis of OFDM based IEEE 802.11g VANET", IJCNWMC, Transteller, volume 3, Issue 1, pp: 47-54, March 2013

[20] Vishal Sharma, Harsukhpreet Singh, Vijay Banga, Mandeep Kaur, "Node-Mobility Sway in IEEE 802.11g MANET", IEEE-ACCT 2013 3rd International Conference on Advanced Computing & Communication Technologies (ACCT 2013), pp:261–266, 2013, India, 10.1109/ACCT.2013.64

[21] Cicconetti, C.; Lenzini, L.; Mingozzi, E.; Eklund, C., "Quality of service support in IEEE 802.16 networks, " *Network, IEEE*, vol.20, no.2, pp.50, 55, March-April 2006 doi: 10.1109/MNET.2006.1607896.

[22] Gronsund, P. and P. Engelstad, “The physical performance and path loss in a fixed WiMAX deployment, ” *Proceeding of International Wireless Communications and Mobile Computing Conference* (IWCMC’07). Honolulu, Hawaii, USA, 2007.

[23] Grondalen O., Gronsund P., Breivik T., Engelstad P., “Fixed WiMAX Field Trial Measurements and Analyses, ” *Mobile and Wireless Communications Summit, 2007. 16th IST*, pp.1–5, 2007 doi: 10.1109/ISTMWC.2007.4299213.

[24] Lee H., Kwon T. and Cho DH., “An Enhanced Uplink scheduling Algorithm based on Voice Activity for VoIP Services in IEEE 802.16d/e System, ” *IEEE communication letters*, Vol. 9, No. 8, August 2005.

[25] Iwan Adhicandra, “Measuring Data and VoIP Traffic in WiMAX Networks, ” Journal Of Telecommunications, Vol. 2, issue 1, pp. 1–6, 2010.

[26] Raman, B.; Chebrolu, K., "Experiences in using WiFi for rural internet in India, " *IEEE Communications Magazine*, vol.45, pp.104-110, doi:10.1109/MCOM.2007.284545, 2007.

[27] M.A. Mohammed, F.W. Zaki and A.M. Elfeki, “Performance Analysis of Scheduling Algorithms for VoIP Services in IEEE 802.16 System, ” *International Journal of Computer Science*, Vol.9, Issue 6, No.3, 2013.

[28] S. Sengupta, M. Chatterjee and S. Ganguly, “Improving Quality of VoIP Streams over WiMAX, ” *IEEE Transactions on Computers*, Vol. 57, No. 2, 2008, pp. 145–156. doi:10.1109/TC.2007.70804.

[29] Ibrahim A. Lawal; Abas Md Said and Abubakar Aminu Mu’azu, “Simulation Model to Improve QoS Performance over Fixed WiMAX using OPNET”, *Research Journal of Applied Sciences, Engineering and Technology* 6(21): 3933-3945, 2013. ISSN: 2040-7459; e-ISSN: 2040-7467, © Maxwell Scientific Organization, 2013.

[30] Grewal, V. and A.K. Sharma, 2010. On performance evaluation of different QoS mechanisms and AMC scheme for an IEEE 802.16 based WiMAX network. *Int. J. Comp. Appl*., 6(7).

[31] Mohammed A. Mohammed; Fayez W. Zaki; and Rania.H. Mosbah, “Improving Quality of VoIP over WiMAX, ” *International Journal of Computer Science and Information Security*, Vol.9, Issue 3, No.3, May 2012

[32] Vikram, M. and Dr. Neena Gupta, 2012. Performance Analysis of QoS Parameters for Wimax Networks. *International Journal of Engineering and Innovative Technology (IJEIT),*

[33] Martufi; G.K.; M. Neves; P.M. Curado; M. Castrucci; P. Simoes; E. Piriand and K. Pentikousis, 2008. Extending WiMAX to new scenarios: Key results on system architecture and test-beds of the WEIRD project. Proceedings of the 2nd European Symposium on Mobile Media Delivery (EUMOB). Oulu, Finland.

[34] R. K. Jha, A. V. Wankhede and U. D. Dalal, “WiMAX System Simulation and Performance Analysis under the Influence of Jamming, ” *International Journal of Scientific Research*, Vol. 1, No. 1, 2010, pp. 20–26.

[35] Rakesh, K., J. Idris, Z. Bholebawa and D. Upena, 2011. Location based performance of WiMAX network for QoS with optimal BSs. *Wirel. Eng. Technol*., 2: 135–145.

Active and Passive Power Conservation Mechanisms in MAC Layer in WSN

Rachna Biala[1] and Jyoteesh Malhotra[2]
[1, 2]CSE Department, Guru Nanak Dev University, Regional Campus, Jalandhar, India
E-mail: [1]rachu_biala@yahoo.com, [2]jyoteesh@gmail.com

***Abstract*—In recent years, WSNs has become an explorative area, particularly due to its potential merits and applications. In order to realize the WSN, many constraints are associated out of which energy conservation is one of them. As the sensor nodes are powered by batteries, which are hardly rechargeable so it is crucial to conserve the power. In this paper, the need of power conservation and factors responsible for power wastage are reviewed from the literature it has been observed that power conservation techniques can be classified as active and passive. The relative issues and merits of these power conservation mechanisms have been extracted and highlighted in this paper.**

Keywords: The MAC, Power Conservation Mechanisms, Issues in MAC layer, Active and Passive Power Conservation

I. Introduction

WSN is becoming centre of topic of research, As WSN have a wide range of Important applications including environment monitoring, medical systems, smart spaces and robotic exploration so Performance analysis and optimization of WSNs, especially its Medium Access Control (MAC) protocols, have received an attention of researchers Conventional MAC protocols for wireless ad hoc networks are designed to maximize throughput and minimize delay.

As sensor nodes are generally battery operated and it is very difficult to replace or recharge the batteries so network lifetime is crucial to WSN. To extend network lifetime, there is a need for efficient power control mechanisms to reduce power consumption in sensor nodes [3]. The power conservation mechanisms are categorized in two parts named as active power conservation mechanisms and passive power conservation mechanisms.

This paper is organized as follows. In section II Active and Passive power conservation mechanisms are described. Section III describes the related work done in power conservation mechanisms. Then comprehensive study of issues is done. Section IV contains open issues and future scope and last section provides the conclusion of the paper.

II. Power Conservation Mechanisms

The main aim of the medium access control (MAC) layer is to ensure that the nodes of network efficiently use the physical medium, while providing error free data transfer to the network layer above it. MAC protocols are influenced by a number of constraints. So in protocol design, there is required a trade-off among several, often contradictory factors, such as quality of service, throughput, and energy efficiency. In wireless sensor networks, the MAC protocol must achieve the following two goals [8]:

1. Create the network infrastructure: Because thousands of sensor nodes are densely deployed, the MAC layer must establish the basic infrastructure needed for hop by hop communication and give the sensor network self-organizing ability.
2. Allow fair and efficient sharing of the wireless communication medium between sensor nodes.

A. Need of Power Conservation

In wireless sensor network, the sensor nodes operates on batteries which have limited life and difficult to recharged so there is need to use energy resources efficiently in order to minimize the energy consumed by the sensor nodes and increase the network lifetime. So, energy efficiency must be considered in every aspect of network design and operation, not only for sensor nodes, but also for the communication of the entire network [14]

B. Factors Responsible for Energy Wastage

The main activities that consume the energy the most are data sensing, data processing, and data communication. Of the three factors, data communication utilizes maximum energy. This involves both data transmission and reception with the ratio between energy consumption in data reception and transmission being 1:2.5 [21].

Other factors that are responsible for energy waste are [1] [2]:

1) Idle listening

It means listening to receive traffic that is not sent. This is especially true in many sensor network applications. If nothing is sensed, the sensor node will be in idle state for most of the time.

2) Collisions

When a transmitted packet is corrupted due to interference, it has to be discarded and require retransmission that increases energy consumption.

3) Overhearing

Since the radio channel is a shared medium, a sensor node may receive packets that are destined to other nodes. Overhearing can be a dominant factor of energy waste when traffic is heavy and sensor nodes are densely deployed.

4) *Packet overhead*

Sending and receiving control packets also consumes energy and less useful data packets can be sent. In applications where only a few bytes of data are transmitted in each message, these overheads can be large.

5) *Traffic fluctuations*

A sensed event will lead to a sudden peak in the sensor network traffic and increase the probability of a collision. The follow on random back off procedure will increase latency and consume energy. When the traffic load approaches the channel capacity, the performance can collapse with little data being delivered while the radio is consuming a lot of energy by repeatedly sensing to identify a clear channel.

C. *Classification of Power Control Techniques in WSN [3]*

1. Active power conservation mechanisms
2. Passive power conservation mechanisms

1) *Active power conservation mechanisms*

Mechanisms that achieve energy conservation by utilizing energy efficient network protocols, rather than turning off the radio (or transceiver) interface of a sensor node.

2) *Passive power conservation mechanism*

Mechanisms that conserve power by turning off the radio (or transceiver) interface

D. *Active Power Conservation Mechanisms*

1) *Multiple ACCESS with Collision Avoidance (MACA)*

It was one of the first channel access protocols proposed to inform other nodes to stay silent before transmitting the data in wireless networks. It also controls power transmission per frame without using carrier sensing. In MACA, a three layer handshake RTS (ready to send)/CTS (clear to send)/DATA is used, which is based on the RTS/CTS exchange. Sender transmits an RTS frame to the receiver to request transmission. If the RTS frame is received by receiver correctly, it will receive the transmission by sending back a CTS frame. When a mobile node overhears some RTS/CTS frames related to the transmissions of other nodes, it is not necessary to remain silent. [4]

2) *Multiple access with collision avoidance wireless*

It is derived from the CSMA/CA protocol, where RTS/CTS/DS/DATA/ACK handshake signaling (or message exchange) are used and merge different back off algorithms. It is a improved version of the MACA protocol, where link layer acknowledgments (ACKs) are added. In the MACAW protocol, a sender can retransmit a packet that was not received by the receiver successfully. The use of the acknowledgment improves the reliability of a wireless link, and consume less energy in transmitting a packet but it does not generally solve the exposed terminal problem and might not normally behave in multicasting. [5]

3) *Power Controlled Multiple Access(PCMA)*

PCMA is a MAC protocol that achieves power controlled transmission and collision avoidance[6]. The goal of the PCMA protocol is to achieve power controlled multiple access within the framework of CSMA/CA base multiple access protocols. At the sender side, the concept of power control with the RTS/CTS based and busy tone based MAC protocols is used.

4) *Power Adaptation for Starvation Avoidance(PASA)*

It is a simple, effective, and independent mechanism with no control message overhead. The PASA protocol adjusts the transmission power in each sensor node to break capture and achieve higher spatial reuse, thus providing fair access of transmission channel to all sensor nodes. Specifically, PASA adjusts the transmission power in each sensor node according to its current condition, so that all mobile nodes in the network can share the medium channel efficiently. [7]

5) *Intelligent medium access with busy tone and power control*

It is another MAC protocol for saving power in mobile ad hoc networks. The main idea of this protocol is to use the exchange of RTS and CTS packets between two intending communicators to determine their relative distance. This information is then utilized to limit the power level on which a mobile host transmits its data packets. The use of lower power can increase channel reuse and thus channel utilization. It can also save the limited battery energy of mobile (or static) sensor nodes and reduces co channel interference with their neighbours. [8]

E. *Passive Power Conservation Mechanisms*

1) *Power Aware Multi access Protocol with Signaling (PAMAS)*

Power aware multi-access protocol with signaling is based on the original MACA protocol with the addition of a separate signaling channel. Saving the battery power at nodes by powering off nodes that are not actively transmitting or receiving packets is its main attribute. Another feature of PAMAS is that it requires two independent radio channels. [9]

It provides solution to overhearing but does not provide solution of idle listening. In this, when mobile sensor node start receiving data frames, it transmits a busy tone to make the other mobile sensor nodes aware when to turn off.

Mobile sensor node will be turned off in following cases [9]:

1. If a mobile sensor node does not have data to transmit and if a neighbour starts transmitting to some other node.
2. If it has data to transmit but at least one of its neighbour pairs is communicating.

A mobile sensor node, which has been turned off when one or more of its neighbour pairs started communication, can determine the time period for which it should be turned off by using a probe protocol. In this probe protocol, the sensor node performs a binary search to determine the time when the current transmission will end. However, the loss of probe frames may cause major power wastage. [9]

2) *Sensor-Medium Access Control(S-MAC)*

It is a sensor MAC layer protocol where sensor nodes are allowed to discover their neighbours and organize a network for communication without requiring the existence of master nodes in the network[11]. It is motivated by PAMAS. S-MAC during transmissions of other nodes, it also sets the radio to sleep . Unlike PAMAS, it only uses in-channel signaling. It reduce contention latency by message passing for sensor-network applications that require store and forward processing as data move through the network. [11]

S-MAC supports multi-hop operation. Its key features are:

1. Periodic listen and sleep
2. Collision avoidance
3. Overhearing avoidance
4. Fixed duty cycle

Sensor MAC (S-MAC) uses three new procedures to decrease energy consumption and support self-configuration. It is a contention-based protocol with low duty cycle. For SMAC, neighbouring nodes of transceiver and receiver are allowed to sleep periodically during transmission which reduces the energy consumption in idle listening . By doing so this scheme put nodes into low duty cycle. S-MAC is based on contention. Periodically sleeping is good in low traffic cases. If a node can sleep for longer time it consumes less energy. Nodes in the S-MAC exchange their sleeping schedule and before going to sleep nodes broadcast their schedule to their neighbours as a SYNC packet. Nodes listen to this sync message and follow it. If they do not get the sync message they make their own schedule.

3) *Timeout-Medium Access Control(T-MAC)*

Timeout-MAC (T-MAC) protocol is an extension to S-MAC protocol in order to improve its performance with respect to latency and throughput under variable traffic load [22]. It uses a timer to indicate the end of the active period instead of relying on a fixed duty cycle schedule. This removes the burden of selecting duty cycle by the applications. It also saves energy by lowering the amount of time spent in idle listening. It is adaptable to traffic conditions changes. The adaptive duty cycle allows T-MAC to automatically adjust to variations in network traffic. But it has the problem of synchronization of the listen periods between nodes within the virtual clusters, due to which early sleeping problem may occur that limits the number of hops a message can travel in each frame time.

F. *Disadvantages of Passive Power Conservation Mechanisms*

As most of the energy is utilized by radio so to conserve energy, it must be turned off whenever not required. But this is not the complete solution as it increases energy consumption rather than decrease. As sensor nodes communicate using short data packets, the data communication energy is dominated by the radio start up energy in most deployments. So, turning the radio off during each idle period will result in negative energy gains. This requires that a well-designed MAC protocol should achieve energy efficiency by finding the right balance between smart radio control and efficient protocol design.

III. RELATED WORK

Syed Jawad Ali and Partha Roy [10](2008) described different protocol for wireless sensor networks and reviewed some proposed MAC protocol like S-MAC to make the network more power efficient. They also discuss benefits and drawbacks of each method. S-MAC has low packet delivery ratio due to adaptive sleeping mechanism, but delay is increased when traffic is high because one node has to wait for others to receive data and latency is also low for periodic sleeping.

Wei Ye, John Heidemann, Deborah Estrin[11] (2008) proposed the S-MAC and analyses the trade-offs between the energy savings and the increased latency due to nodes sleep schedules by comparing S-MAC with protocols that do not have periodic sleep such as the IEEE 802.11, for a packet moving through a multi-hop network. They have used the motes and TinyOS platform to test the S-MAC and concluded that the S-MAC has the ability to make trade-offs between energy and latency according to traffic conditions.

Liqiang Zhao, Le Guo, Li Cong, Hailin Zhang [12] (2009) modeled the problem of energy-efficient MAC protocols in WSNs as a game-theoretic constraint optimization with multiple objectives. They provided an auto digressive back off mechanism to implement the game in current WLANs.

N. Dimokas, D. Katsaros, Y. Manolopoulos [13](2010) considered the problem of energy efficiency and multihop communication and proposed an energy-efficient distributed clustering protocol for wireless sensor networks called called GESC (GEodesic Sensor Clustering) that was based on a metric for characterizing the significance of a node. The protocol achieves small communication complexity and linear computation complexity. Experimental results attest that the protocol improves network longevity.

Suraiya Tarannum [14] (2010) discusses the WSNs, their characteristics, issues, Energy Conservation Challenges in Communication Protocols and Design Issues in WSNs and applications. The energy conservation challenges and related issues emphasize the need for energy saving and optimizing protocols to increase the lifetime of sensor networks

Ines Slama, Badii Jouaber, Djamal Zeghlache [15] (2010) introduced an adaptive hybrid medium access protocol called I-MAC for wireless sensor networks .IMAC uses adaptive prioritization mechanism in order to have more chances of accessing the radio resources by the sensor node and to reduce the chances of collision. They also evaluated the performance of IMAC through simulations over NS2 and concluded the improvement, compared to Z-MAC, mainly in energy efficiency, channel utilization, loss ratio and delay.

Wen-Hwa Liao, Hung-Chun Yang [16] (2012) provide a data storage scheme that supports energy-efficient mechanism and scheme was based on grid based architecture. They evaluated the network lifetime and residual power percentage of sensors by varying the number of sensor nodes, storage events and concluded that both network lifetime and residual power of their scheme increases with the increase in the number of sensors and decreases with increase in data storage events.

Amir Esmailpour, Moataz Alfaraj, Jamal Alfaraj, and Gelareh Kokabian [17] (2013) Proposed solar rechargeable system that checks the radiation level in the area and dynamically adjusts the recharging cycle of the sensor node battery based on the level of radiation .They designed two methods where they calculated electrical energy and battery lifetime with 11%, 15%, 20% and 25% cell efficiencies in two different areas having different solar radiation environment and concluded that battery lifetime is longer in stronger radiation area than in weak radiation area and battery lifetime varies diversely by adjusting PV efficiencies and radiation levels within the same area as well as in different areas.

Dharam Vir, S.K. Agarwal, S.A. Imam [18] (2013) discuss two mechanisms that effect energy consumption: power control and power management. They analysis power control mechanisms of MAC Protocol for wireless sensor network using QualNet simulator and made modifications in the virtual carrier sensing scheme of 802.11 MAC in order to reduce the power consumption and to increase total throughput.

Abayomi M. Ajofoyinbo[19](2013) proposed the novel energy-efficient MAC protocol based on the use of duration value in transmitted packets to setup varying sleep/wake-up schedules for neighbouring nodes of the receiver. The effectiveness of this proposed Packet Duration Value based MAC (PDV-MAC) protocol was tested via Simulation implemented in Visual C# and MATLAB.

Mahir Meghji, Daryoush Habibi [20] (2014) extended the previous research by investigating transmission power control(TPC) in single hop and multi hop WSN using typical Telosb platform to describe the benefits of single hop TCP over multihop TPC and concluded that transmitting in single-hop networks at lower transmission power levels reduced energy consumption by up to 23 %while maintaining reliability.

TABLE 1 COMPREHENSIVE SUMMARY OF ISSUES

Paper	Issues	Remarks
Energy Efficient MAC Protocols for Wireless Sensor Networks [11](2008)	To analyses the trade-offs between the energy savings and the increased latency due to nodes sleep schedules.	Researchers reported that Future work can be done in system scaling studies and parameter analysis. More tests can be done on larger test beds with different number of nodes and system complexity.
An Energy-Efficient MAC Protocol for WSNs: Game Theoretic Constraint Optimization with Multiple Objectives [12](2009)	To provide simple method to address the sleeping probability in WSN.	Researchers concluded that based on G-Conopt, each sensor node can achieve optimal performance independently under limited energy consumption.
Energy-efficient distributed clustering in wireless sensor networks [13](2010)	Issues of network node clustering.	GESC is very efficient clustering protocol and it is able to show significant performance gains in terms of communication cost (few transmitted messages) and also in terms of network longevity.
Energy Conservation Challenges in Wireless Sensor Networks: A Comprehensive Study [14](2010)	1. Energy control 2. to satisfy the QoS parameters	The energy conservation challenges and related issues emphasize the need for energy saving and optimizing protocols to increase the lifetime of sensor networks.

Table 1 (Contd.)…

…Table 1 (Comprehensive Summary of Issues)

Paper	Issues	Remarks
Priority Based Hybrid MAC for Energy Efficiency in Wireless Sensor Networks [15](2010)	To develop novel communication protocols and algorithms that efficiently tackle the resource constraints and application requirements	Researchers look forward to achieve further enhancements over I-MAC through the proposal of a more dynamic and efficient mechanism for priority adaptation and to implement it over TinyOS to validate the simulation results they provided.
An Energy-Efficient Data Storage Scheme in Wireless Sensor Networks [16](2012)	As Data storage schemes cannot perform well based on energy-efficient protocols. So, the issue was to propose a data storage scheme to support energy-efficient mechanism.	Both network lifetime and residual power of scheme increases as the number of sensors increases and decreases with increase in data storage events.
Energy Conservation for Wireless Sensor Networks Using Solar Rechargeable Power Source [17](2013)	To improve the power Conservation of a sensor node by adjusting the recharging cycle of the solar-fed batteries based on the geographical areas and radiation levels of the areas and to increase the battery lifetime.	Researchers plan to study and try various types of batteries and adjust different environmental conditions such as temperature and humidity and see their effect on the battery lifetime. they also expect different newly proposed Cluster-based protocols and algorithms to be investigated and compared with other approaches to save the energy in WSN
Analysis of Power Control Mechanisms of MAC Protocol for wireless Sensor Networks [18](2013)	To analysis a Power Control MAC protocol for WSN with overall power consumption and improve the throughput of the network.	Judge different mobility modes of the nodes to make it more suitable for mobile ad hoc networks.

IV. Open Issues and Future Scope

Based on the extensive literature survey done in the previous section, the major issues have been extracted that includes eliminate collision, idle listening in order to conserve energy and sensor network lifetime which need immediate attention of the research community. Another important issue is to implement different types of batteries in different environmental conditions so as to check the effect on the battery lifetime. Also it requires focus on other aspects in wireless sensor network like latency, throughput, cross layer protocol design etc.

V. Conclusion

WSN has many issues but power conservation is one of the most important issues. Here various active and passive power conservation MAC protocol for wireless sensor networks to make the network more power efficient are discussed. We also consider various issues described by the researchers. Different applications of WSN have different requirements like in environment monitoring application such as checking pressure, temperature, and humidity etc, where the use of power is more sensitive to make network more energy efficient. It is hoped that survey done here in this paper will prove to be helpful to researchers working in the area of WSN in general and power conservation in particular.

Acknowledgment

We would like to thank almighty for his constant blessings. Then we would like to thank our families and friends for helping and supporting us throughout the making of this paper.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E.Cayirci.Wireless sensor networks: A survey. Computer Networks, 38(4):393–422, March 2002

[2] A Survey of MAC protocols for wireless sensor network, Rajesh Yadav, Shirshu Varma, N. Malaviya, UbiCC Journal, Volume 4, Number 3, August 2009

[3] C. Srisathapornphat and C. - C. Shen, “ Coordinated power conservation for ad hoc networks ”, in Proceedings of 2002 IEEE International Conference on Communications (ICC ’ 02), vol.5, New York, Apr.-May 2002, pp. 3330 – 3335

[4] P. Karn, “ MACA A new channel access method for packet radio ”, in Proceedings of the ARRL CRRL Amateur Radio 9th Computer Networking Conference, Redondo Beach, CA, Apr. 1990, pp. 134–140 .

[5] V. Bharghavan, A. Demers, S. Shenkar, and L. Zhang, “ MACAW: A media access protocol for wireless LANs ”, in Proceedings of ACM SIGCOMM ’ 94, London, UK, Sept. 1994, pp. 212– 225 .

[6] J. Monks, V. Bharghavan, and W. - M. Hwu, “ A power controlled multiple access protocol for wireless packet networks ”, in Proceedings of IEEE INFOCOM ’ 01, vol. 1, Anchorage, AK, Apr. 2001, pp. 219 – 228.

[7] J. Chen, S. - H. G. Chan, Q. Zhang, W. - W. Zhu, and G. Chen, “ PASA: Power adaptation for starvation avoidance to deliver wireless multimedia ”, IEEE Journal on Selected Areas in Communications, vol. 21, no. 10, Dec. 2003, pp. 1663 – 1673

[8] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey, Computer Networks, 38(4):393–422, March 2002

[9] S. Singh and C. Raghavendra, “ PAMAS: Power aware multi - access protocol with signaling for ad - hoc networks ”, ACM SIGCOMM Computer Communication Review, vol. 28, no. 3, July 1998, pp. 5 – 26 .

[10] Syed Jawad Ali and Partha Roy, ” Energy Saving Methods in Wireless Sensor Networks” Technical report, IDE0814, May 2008.

[11] Wei Ye, John Heidemann, Deborah Estrin, An Energy-Efficient MAC Protocol for Wireless Sensor Networks”pages-1-69, 2008.

[12] Liqiang Zhao, Le Guo, Li Cong, Hailin Zhang, ” An Energy-Efficient MAC Protocol for WSNs: Game Theoretic Constraint Optimization with Multiple Objectives”pages-358–364, June 10, 2009.

[13] N. Dimokas, D. Katsaros, Y. Manolopoulos, ” Energy-efficient distributed clustering in wireless sensor networks ”Journal of parallel and distributed computing, pages 371-383, 2010.

[14] Suraiya Tarannum, “Energy Conservation Challenges in Wireless Sensor Networks: A Comprehensive Study”, 2010.

[15] Ines Slama, Badii Jouaber, Djamal Zeghlache, ” Priority Based Hybrid MAC for Energy Efficiency in Wireless Sensor Networks”, pages-755-767, September 15, 2010

[16] Wen-Hwa Liao, Hung-Chun Yang, ” An Energy-Efficient Data Storage Scheme in Wireless Sensor Networks”, IEEE Network Operations and Management Symposium (NOMS): Short Papers, 2012

[17] Amir Esmailpour, Moataz Alfaraj, Jamal Alfaraj, and Gelareh Kokabian, ” Energy Conservation for Wireless Sensor Networks Using Solar Rechargeable Power Source” ASEE Northeast Section Conference Reviewed Paper Norwich University March 14-16, 2013.

[18] Dharam Vir, S. K. Agarwal, S. A. Imam, ” Analysis of Power Control Mechanisms of MAC Protocol for wireless Sensor Networks” International Journal of Science and Research (IJSR), Volume 2 Issue 3, March 2013.

[19] Abayomi M. Ajofoyinbo, ” Energy Efficient Packet-Duration-Value Based MAC Protocol for Wireless Sensor Networks” September 14, 2013.

[20] Mahir Meghji, Daryoush Habibi, ” Investigating transmission power control for wireless sensor networks based on 802.15.4 specifications “Telecommunications Systems Volume 56**,** Issue 2, June 2014, 299-310, june 2014

[21] R. Doss, G. Li, V. Mak, S. Yu, and M. Chowdhury. The crossroads approach to information discovery in WSNs. Lecture Notes in Computer Science 4094, January 2008.

[22] T. van Dam and K. Langendoen. An adaptive energy-efficient MAC protocol for wireless sensor networks. In Proceedings of the International Conference on Embedded Networked Sensor Systems (SenSys), pages 171–180, November 2003

A Comprehensive Survey on Various Evolutionary Algorithms on GPU

Satvir Singh[1], Jaspreet Kaur[2] and Rashmi Sharan Sinha[3]
[1,2,]*Department of Electronics & Comm. Engineering,*
SBS State Technical Campus, Moga Road Ferozepur–152004, Punjab
E-mail: [1]drsatvir.in@gmail.com, [2]er.jaspreetkaur1@hotmail.com,
[3]sinharashmisinha@hotmail.com

***Abstract*—This paper presents a comprehensive survey on parallelizing computations involved in optimization problem on Graphics Processing Unit (GPU) using CUDA (Compute Unified Design Architecture). GPU have multithread cores with high memory bandwidth which allow for greater ease of use and also more radially support a layer body of applications. Many researchers have reported significant speedups with General Purpose computing on GPU (GPGPU). Stochastic meta-heuristic search algorithms, e.g., Mixed Integer Non-Linear Programming (MINLP), Central Force Optimization(CFO), Genetic Algorithms (GA), and Particle Swarm Optimization(PSO), etc. are being investigated nowadays for improved performance with processing power of GPU. From study it is found that GPGPU shows tremendous speedups from 7 times in Steady State GAs to 10, 000 times speedups in CFO.**

Keywords: GPU, GPGPU, CUDA, MINLP, PGMOEA, CGA, CFO, Optimization Algorithms

I. INTRODUCTION

General Purpose GPU Computing really took off when CUDA and Stream arrived in late 2006 [1]. GPU constitutea tremendous step towards a usable, suitable, scalable and manageable future-proof programming model [2]. Optimization works are significantly parallel, and so GPUs evolved as large-scale general purpose computation machines [3] [4] [5] [6]. With the advent and large availability of General Purpose Graphics Processing Units and the development and straightforward applicability of the Compute Unified Device Architecture platform, several applications are being benefited by the reduction of the computing time [7]. GPGPU-based architecture, aiming at improving the performance of computationally demanding optimizations for identifiable specific mapping parameters, one can reduce total execution time drastically and also, improve greatly the optimization process convergence. Application performance can be significantly improved by applying memory access pattern-aware optimizations that can exploit knowledge of the characteristics of each access pattern [3]. To evaluate the effectiveness of our methodology, we have created a tool that incorporates our proposed algorithmic optimizations and report on execution speedup using selected benchmark kernels that cover a wide range of memory access patterns commonly found in GPGPU workloads [8]. Graphics Processing Units (GPUs) are widely used among developers and researchers as accelerators for applications outside the domain of traditional computer graphics. In particular, GPUs have become a viable parallel accelerator for scientific computing with low investment in the necessary hardware.

In this paper, various sections describe different Evolutionary Algorithms (EAs) and their gained efficient speedup on GPU using CUDA. Sections II to V are dedicated to various GAs variants those have been investigated by various researchers. Sections VI to X presents Cellular automata, CFP, Multi-objective Optimization, PSO Differential Evolution (DE).Finally Section XI and XII present study on GAs Versus DE and conclusions at the end.

II. ISLAND BASED GENETIC ALGORITHM

Island based genetic algorithm (GA) is implemented on Multi-GPU in [27], for solving the Knapsack problem. The main motiveis to speed up the GA by using a cluster of NVIDIA GPU and comparing the execution time of single GPU with a multicore CPU.

The population of proposed GA is organized in two one-dimensional arrays. First array representing the genotype and the other array represents the fitness value. The GA parameter such as population and chromosome size, the crossover and mutation rates, the statistic collection and migration interval, the total number of evaluated generations etc. are filled with command line parameters, this maintained structure is stored at GPU constant memory [28]. The basic concept to maximize GPU utilization is to control thread divergence and amalgamate all memory accesses using this algorithm [28]. Firstly, a hash function generator based stateless random number is generated [29][30]. Then, the genetic material is exchanged of two parents using crossover and mutation process by performing the binary tournament selection to create a new individual. As the new offspring is created fitness evaluation is carried out. Next, the parent is replaced by the offspring with the help of entire warp if the fitness value of latter is higher than the former. The good individuals are migrated from the adjacent lower index island to the higher index island which is arranged in the unidirectional ring topology. Lastly, all the statistical data from the local island and from the global gathering process are collected [31].

The analysis illustrates that as the individual per GPU and number of islands increases the fitness value increases. Secondly, the execution time is invariant for island size up to 512 and then elevate linearly beyond 512. All in all, the implemented Island based GA leads to the GPU performance of 5.67 TFLOPS.

III. Advanced Genetic Algorithm

With the increasing advent of GPGPU using CUDA, the stochastic algorithm of advanced Genetic Algorithm is used to solve non-convex MINLP and non-convex Non-linear Programming (NLP) problems [9]. MINLP refers to mathematical programming algorithms that can optimize both continuous and integer variables, in a context of nonlinearities in the objective function and/or constraints. MINLP problems involve the simultaneous optimization of discrete and continuous variables. These problems often arise where one is trying to simultaneously optimize the system structure and parameters. This is difficult because optimal topology is dependent upon parameter levels and vice versa [9].

In many design optimization problems, the structural topology influences the optimal parameter settings so a simple de-coupling approach does not work: it is often not possible to isolate these and optimize each separately. Finally, the complexity of these problems depends upon the form of the objective function. In the past, solution techniques typically depended upon objective functions that were single-attribute and linear (i.e., minimize cost). However, real problems often require multi-attribute objectives such as minimizing costs while maximizing safety and/or reliability, ensuring feasibility, and meeting scheduled deadlines. In these cases, the goal is to optimize over a set of performance indices which may be combined in a nonlinear objective function. For efficient utilization of GPU parallel resources adaptive resolution genetic algorithm (arGA) is developed [9].Through this algorithm the intensity of each individual is beamed using entropy measures. The algorithm is tested for different benchmarking problems [9] having different levels of difficulty. Parallelization of arGA and the arLS (local search) operators is done to gaina significant speedup. The results of the tests shows a speedup of 42x with single precision and 20x with double precision overnVidia Fermi C2050 GPU [9].

IV. Steady State Genetic Algorithm

Steady-State GA is implemented on GPU using CUDA in [23], where population individual data is accessed parallel to effectively speedup the process. The optimization problem is effectively solved by the means of Evolutionary Computing [24]. The steady state Genetic Algorithm is used to access optimization algorithms. These algorithms basically have selection for the reproduction and selection of survival implementation with concurrent kernel execution [22].

The implementation of Steady-State GA is done as follows: Firstly, from the population two individual (parents) are received by Streaming Multiprocessors (SM). Then, the kernel generates random number as GAs are stochastic search processes. BLX-α is adopted as blend crossover in the crossover process. The crossover operation is executed with two parents in SM, and the two offspring yielded are stored in the shared memory. Next, uniform mutation, fitness based sorting process and selection process is executed. This whole process is repeated until the loop terminates. Four test functions of the optimization problem Hyper sphere, Rosen rock, Ackley, and Griewank were used for comparing GPU and CPU computation on implementing Steady-State GA. The study is first performed upon CPU then with nVidia GeForce GTX480GPU gives a speedup of 3x to 6x then the previous implementation on CPU [19]. Moreover, the speed up ratio for Generational GA is much better than Steady-State GA on GPU since computational granularity is very small in the latter. So a large amount of execution time is occupied by the latency caused by the kernel calls. However, in terms of function values Steady-State GA are more efficient.

V. Cellular Genetic Algorithm

Genetic Algorithm have a subclass known as Cellular Genetic Algorithm (cGA) which provides the data of population structured in several specified topologies [19]. Cellular Genetic Algorithm (cGA) is implemented for multi-GPU to accelerate the execution process so that the system could be more efficient. cGAis used because of its high performance and swarm intelligence structure. Until the global optimum region is reached, cGA is able of keeping a high diversity in population.

To manage the multi-GPU utilization each CPU thread is held responsible for one GPU device which is known as multi-threaded mode. Firstly, each CPU thread is associated to one GPU. This can be done if common structure (toroidal grid) is designed for all CPU threads. Then, the population is divided into subpopulations which is stored in the global memory of each GPU. Each GPU works individually irrespective of other GPU and the process is same as performed with single GPU implementation. To ensure that every GPU had finished its work a synchronization barrier is used and lastly, data is collected and transferred to other GPUs. The process executes until the while loop in pseudo code of cGA terminates.

Three discrete optimization problems: Colville Minimization, Error Correcting Codes Design Problem (ECC) and Massively Multimodal Deceptive Problem (MMDP), and three continuous ones, Shifted Griewank function, Shifted Restringing function and Shifted Rosen rock function [20] [21] were selected for comparing the algorithm in terms of efficiency and efficacy. Statistical tests [19] are performed for each problem to ensure that the results are statistically significant. A common parameter, population size is used to make a meaningful comparison among all the algorithms.

The analysis shows the average speed up with respect to CPU version ranges from 8 to 771 and for single GPU it is alike multi-GPU, with a little overhead in the latter case. The multi-GPUis more prominent in paralleling the algorithm and producing accurate results as there is a need of special maintenance to perform same experiment upon single GPU.

Genetic Algorithm is tested and evaluated on parallel implementation on C-CUDA API on the parameters like population size, number of threads, problem size and problem of differing complexities with variation in the population individuals [19]. For an efficient implementation on GPGPU the solution is thoroughly implemented along with the operators like random number generation, initialization, selection operation, and mutation operations [13]. The nVidia GeForce 8800GTX shows overall speedup of 40–400 on three different test problems [22]. Thus parallel implementation is more effective then sequential process as compared with clock time and accuracy.

VI. Cellular Automata

Cellular Automata have various real life applications like physical system modeling, road traffic simulation, artificial life simulation, etc. [14] [15] [16]. Cellular automata design evolved from evolutionary algorithm and a part of Genetic Algorithm which is complex in nature.

The Algorithm is parallelized and implemented upon GPGPU shows an efficient reduction in execution time. The rules of Cellular Automata take longer time period in evolution in sequential execution. The same Genetic Algorithm shows 31.34x to 314.94x speedup when executed upon nVidia GeForce FX280 GPU which is a significant reduction in execution time [17].

VII. Central Force Optimization

The metaheurestic algorithm CFO is implemented upon GPGPU using local neighborhood and implemented CFO concepts [25]. The calculation of CFO independent upon the movement of probes which are scattered allover the space. The probes then slowly move towards the probehaving highest mass or fitness. PR-CFO is the most evaluated algorithm with the measures of initial position and acceleration vectors, fitness evaluation and probe movements [26]. The test problems are having the dimension of 30 to 100 of four different examples of Pseudo random CFO (PR-CFO). The PR-CFO is tested with four test types i.e. Ring, Standard, CUDA, CUDARing. PR-CFO shows a speedup of 4 to 400 using CUDA. PRCFOring and PR-CFO CUDA ring on nVidia Tesla C1060 shows10, 000 times faster results as compared with standard PR-CFOalgorithm [26].

VIII. PGMOEA

The general Purpose GPU is efficiently used in optimizing the multiple objective problems. The particle gradient Multi-objective Evolutionary Algorithm (PGMOEA) is used to solve optimization problems. PGMOEA is first experimented on CPU and then after parallelizing the algorithm executed upon GPU which formed agreat speedup results [18]. The first step to implement PGMOEA is to read parameters such as population size, dimension size, maximum iterative generations, crossover rate, mutation rate and initialize particle texture array. Blank texture array i.e objective, rank value, entropy and free energy array are then generated to store different results. Next, the particle texture array is is loaded to GPU to calculate the rank of all the particles and the results are then stored in rank value texture array. The particles are sorted in the decreasing order of their ranks to make a mating pool. The higher order rank value particles are selected to perform crossover and mutation operation using Guo’s algorithm. The new particles generated through this process are then replaced with the last particles which have lower rank in mating pool to get a new population. The program is terminated if the halt condition is satisfied else particle texture array is again loaded to GPU and the process is repeated again.

Table I Speedup Comparison (Source [18])

PGMOE Algorithm	Example 1		Example 2	
	Time (s)	Speedup	Time (s)	Speedup
On GPU	0.97	9.95	0.83	10.64
On CPU	9.01	1.04	8.02	1.10

The experiment is conducted upon two different examples. The first example shows a speedup of 9x with nVidia GeForceGTX285 then CPU result, while the second example is 10x faster than that of CPU [18]. The speedup comparison is shown bellowing Table I.

IX. PSO Algorithm

PSO is a met heuristic algorithm works by having a swarm of particles [10]. These particles are moved around in the search-space according to a few simple formulae. The movements of the particles are guided by their own best known position in the search-space as well as the entire swarm best known position [11].

When improved positions are being discovered these will then come to guide the movements of the swarm. PSO is one of the types of Evolutionary Algorithm used to optimize the multiple objective problems. When an optimization problem involves more than one objective function, the task of finding one or more optimal solutions is known asmulti-objective optimization [10].

For implementing PSO code in C-CUDA the allocation of vector/matrix is done on the device.

Random numbers are generated using Mersenn Twister code and then based on objective functions are evaluated. After evaluation the global best particle of whole swarm is updated. Next, the sum and multiplication operations are performed on the vectors which describe the particle. The benchmark functions with many local minima mentioned in table A1 appendix [12] are selected. The algorithm is tested upon three different platforms of C, Matlab and C-CUDA. The parallel implementation of PSO on nVidia GTX 280 gives 17 to 41 times speedup in computing time in C-CUDA as compared with the Cand Matlab as Shown in Fig. 1 [12].

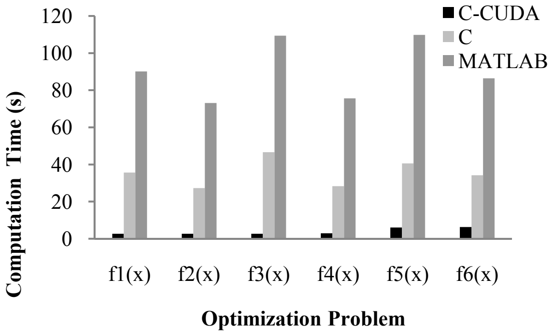


Fig. 1 Computing Time for C-CUDA, C, and MATLAB (Source [12])

TABLE II RUNNING TIME RESULTS USING C-CUDA AND C FOR THE BENCHMARK OPTIMIZATION PROBLEMS WITH 100 DIMENSIONS, 100 INDIVIDUALS, 10, 000 AND 100, 000 ITERATIONS (SOURCE [13])

Benchmark Functions	Implementation Language	10, 000 Iterations			100, 000 Iterations		
		Computing Time(s)	Standard Deviation	Speedup	Computing Time(s)	Standard Deviation	Speedup
Schwefel-F_1(x)	C-CUDA	0.64	0.01	NA	27.72	0.45	NA
	C	9.59	0.21	15.05	983.65	30.91	35.48
Rastrigin-F_2(x)	C-CUDA	0.64	0.01	NA	27.47	0.38	NA
	C	8.39	0.27	13.15	900.97	29.50	32.80
Ackley-F_3(x)	C-CUDA	0.72	0.01	NA	36.33	1.38	NA
	C	7.10	0.25	9.91	736.68	39.56	20.28
Griewank- F_4(x)	C-CUDA	0.69	0.01	NA	31.44	0.37	NA
	C	10.39	0.44	14.96	1005.35	4.33	31.98
Generalized penalized function-F_5(x)	C-CUDA	0.76	0.02	NA	37.88	1.16	NA
	C	13.75	1.13	18.07	1344.59	72.33	35.50
Generalized penalized function-F_6(x)	C-CUDA	0.72	0.01	NA	37.44	1.50	NA
	C	13.76	1.36	19.04	1300.11	81.40	34.73

TABLE III COMPARISON TABLE OF DIFFERENT EVOLUTIONARY ALGORITHMS ON GPU AND CPU

Algorithm		Experimental Set up		Time		Speed up	
		GPU (nVidia)	CPU	GPU	CPU	GPU	CPU
Island based GA		GTX580	Intel Xeon Six-Core	5.67 TFLOPS	–	653.68	11.32
Advanced GA		C2050 (Double precision)	Intel Core 2 Duo	–	–	20x	
		C2050 (Single precision)	Intel Corei7			40x	
Steady-state GA		Geforce GTX480	Intel Core i7	4.874 - 4.780s	14.46-28.56s	3.0x-6.0x	
Cellular GA		GTX-285	Intel Quad processor	0.021 - 1.821s	0.266 - 1450.415s	8 - 771	–
Binary and Real coded GA		Tesla C1060	AMD Athlon 64 X2 Dual Core	RGA 0.003365s - 22.534169s	RGA 0.071639 - 4851.69		RGA 21.289 - 215.304
Cellular Automata	Laptop	GeForce 8600M GS	Intel Core 2 Duo Processor	–	–	31.34x	
	Work-station	GeForce FX 280	Intel Core 2 Duo Processor	–	–	314.97x	
PGMEOA		GeForce GTX285	Intel Core (TM)2 Q8200	0.83s - 0.97s	8.02s - 9.01s	9.95 - 10.64	1.04 - 1.10
CFO		Tesla C1060	Intel Xeon 5504 Quad Core	–	–	10, 000x	

Table III (Contd.)…

…Table III (Comparison Table of Different Evolutionary Algorithms on GPU and CPU)

Algorithm	**Experimental Set up**		**Time**				**Speed up**	
	GPU (nVidia)	CPU	GPU		CPU		GPU	CPU
PSO	GTX 280	AMD Athlon x2 3800 + 2.0 GHz Dual Core	–		–		17 to 41x	
DE	GTX285	AMD Athlon x2 5200 + 2.7 GHz Dual Core	$F_1(x)$ 5.84	$F_6(x)$7.25	$F_1(x)$481.38	$F_6(x)$682.76	20x to 35x	
Many threaded DE and GA	Tesla C2050	Dual Core AMD Opteron	–		–		DE 19 - 34 x	

X. Differential Evolutionary Algorithm

GPGPU is proved to be great architectural unit in reducing the processing time [13]. The Differential Algorithm which is one of the parts of EAs is implemented upon CPU using C-CUDA. The motivating features of Differential Algorithm are easy for parallelization and convergence properties which intern gives an appropriate result. The algorithm is first tested upon CPU then on nVidia GTX285 with 1GB GDDR3 GPU with the speedup outcomes.

The implementations of DE algorithm and benchmark functions are same as used for PSO implementation on GPU. GPU gives 20x to 35x faster results which proves GPU is much more effective and efficient than Differential Algorithm on CPU [13]. The Speedup comparison results are shown in Table II.

XI. DE Versus GAs

In this paper [32], two evolutionary meta-heuristic algorithms DE and GAs, many threaded implementation is done on CUDA and results were compared when Independent task scheduling is solved. Mapping of set of task to a set of resources is known as Independent task scheduling [33] [34]. Since it is a NP-complete problem so two objectives make span and flow time are used during task mapping for optimization.

Whole meta task can be accelerated by minimizing make span and the efficient utilization of the computing environment can be done by minimizing flow time.

$$makespan = \min_{S \epsilon Sched}\{\max_{j \epsilon Jobs} F_j\} \quad (1)$$

$$flowtime = \ min_{S \epsilon Sched}\{\textstyle\sum_{j \epsilon Jobs} F_j\} \quad (2)$$

Real coordinates are used in DE [35]for encoding real vectors. Truncation of real encoded vector coordinates is done to translateit into schedule representation. For this fitness function f(S):Sched→R is defined which evaluates each schedule

$$f(S) = \lambda makespan(S) + \ (1 - \lambda)\frac{flowtime(S)}{m} \quad (3)$$

For minimization purpose a standard proposed in [34] is used. The simulation matrix is derived from ETC matrix. The time taken by the machine to execute a task can be estimated using Expected Time to Compute (ETC) matrix. The average final fitness value is calculated by both the algorithms for each ETCmatrix. The analysis shows that the DE is better than GA for solving Independent task scheduling problem and leads to better results for many threaded implementation on CUDA.

XII. Conclusion

In this paper we present different optimization algorithm with tremendous speedups in the computation time. The overall GPU performance of multi-GPU Island-based GA for solving Knapsack problem reaches 5.67 TFLOPS. MINLP archived an overall speedup of 20x to 42x using nVidia Tesla C2050 GPU as compared to Intel Core i7 920 CPU processor. On implementing Steady state GA on a GPU approximately 6 times faster results are obtained than the corresponding CPU implementation. The implementation of Cellular Genetic Algorithm for a multi-GPU platform leads to speedup range from 8 to 771 with respect to the CPU version. The new binary-coded and real-coded Genetic Algorithm using CUDA leads to a performance improvement with the speedup of 40x to 400x. Central Force Optimization (CFO)results in reduction of computing time and a speedup of 10, 000x. The computing time gets accelerated up to 17 to 41 times in C-CUD Aafter PSO is paralleled implemented on NVIDIA GTX280. GPU is much more effective and efficient for Differential Evolutionary algorithm since it gives 20x to 35x faster results than CPU. The Cellular Automata shows 314.97x as compared with the sequential implements. The advantage of GPU computing is that it is fast and cheap. A theoretical 1.5 TFLOPS is obtained by the newest nVIDIA GTX 580 at $500. The major drawback is not all algorithms can have theoretical speedup and are hard to program.

References

[1] K. S. Perumalla, “Discrete-event Execution Alternatives on General Purpose Graphical Processing Units (GPGPU), ” in *Proceedings of the 20th Workshop on Principles of Advanced and Distributed Simulation*, 2006, pp. 74–81.

[2] J. A. Jablin, P. McCormick, and M. Herlihy, “Scout: high-performanceheterogeneous computing made simple, ” in *IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW), 2011*, 2011, pp. 2093–2096.

[3] Bustamam, K. Burrage, and N. A. Hamilton, “Fast Parallel Markov Clustering in Bioinformatics using Massively Parallel Graphics Processing Unit Computing, ” in *2010 Ninth International Workshop on Parallel and Distributed Methods in Verification and Second International Workshop on High Performance Computational Systems Biology*, 2010, pp. 116–125.

[4] C. Xue-bin *et al.*, “Data Processing in Space Weather Physics Models in the Meridian Project, ” in *2010 Ninth International Symposium on Distributed Computing and Applications to Business Engineering and Science (DCABES)*, 2010, pp. 342–345.

[5] F. Pei-qin, D. Liang-Long, L. Xiao-Ting, and J. Chao-bo, “Design and Implementation of Remote Parallel Computing System based on Multi-Platform, ” in *2010 International Conference on Internet Technology and Applications*, 2010, pp. 1–4.

[6] M. Al Hajj Hassan and M. Bamha, “An Efficient Parallel Algorithm for Evaluating Join Queries on Heterogeneous Distributed Systems, ” in *2009 International Conference on High Performance Computing (HiPC),* IEEE, 2009, pp. 350–358.

[7] D. Luebke, M. Harris, N. Govindaraju, A. Lefohn, M. Houston, J. Owens, M. Segal, M. Papakipos, and I. Buck, “GPGPU: GeneralPurpose Computation on Graphics Hardware, ” in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, 2006, p.208.

[8] W. B. Langdon and M. Harman, “Evolving a CUDA Kernel from an nVidiaTemplate, ” in *2010 IEEE Congress on Evolutionary Computation (CEC)*, 2010, pp. 1–8.

[9] Munawar, M. Wahib, M. Munetomo, and K. Akama, “Advanced Genetic Algorithm to Solve MINLPProblems over GPU, ” in*2011 IEEE Congress on Evolutionary Computation (CEC)*, 2011, pp. 318–325.

[10] J. Kennedy, J. F. Kennedy, and R. C. Eberhart, *Swarm Intelligence.* Morgan Kaufmann, 2001.

[11] J. Kennedy and R. Mendes, “Population Structure and Particle Swarm Performance, ” *Proceedings of the World on Congress on Computational Intelligence*, vol.2, pp. 1671–1676, 2002.

[12] L. de P Veronese and R. A. Krohling, “Swarm’s Flight: Accelerating the Particles using c-CUDA, ” in *IEEE Congress onEvolutionary Computation*, 2009, pp. 3264–3270.

[13] L. De Veronese and R. A. Krohling, “Differential Evolution Algorithm on the GPU with c-CUDA, ” in *2010 IEEE Congress onEvolutionary Computation (CEC)*, 2010, pp. 1–7.

[14] L. J. Durbeck and N. J. Macias, “The Cell Matrix: An Architecture for Nanocomputing, ” *Nanotechnology*, vol. 12, no. 3, p. 217, 2001.

[15] M. Gardner, “Mathematical Games: The Fantastic Combinations of John ConwaysNew Solitaire Game Life, ” *Scientific American*, vol. 223, no.4, pp. 120–123, 1970.

[16] M. Tomassini, M. Sipper, and M. Perrenoud, “On the Generation of High-quality Random Numbers by Two-dimensional Cellular Automata, ”*IEEE Transactions onComputers*, vol. 49, no. 10, pp. 1146–1151, 2000.

[17] L. Zaloudek, L. Sekanina, and V. Simek, “GPU Accelerators for Evolvable Cellular Automata, ” in *2009Computation World:Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns*, 2009, pp. 533–537.

[18] X. Yue, Z. Wu, and K. Li, “Particle Gradient Multi-Objective Evolutionary Algorithm based on GPU with CUDA, ” in *2010 International Symposium onInformation Science and Engineering (ISISE)*, 2010, pp. 540–544.

[19] P. Vidal and E. Alba, “A Multi-GPU Implementation of a CellularGenetic Algorithm, ” in *2010 IEEE Congress onEvolutionary Computation (CEC)*, 2010, pp. 1–7.

[20] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y.-P. Chen, A. Auger, and S. Tiwari, “Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization, ” *KanGAL Report*, vol. 2005005, 2005.

[21] K. Tang, X. Y´ao, P. N. Suganthan, C. MacNish, Y.-P. Chen, C.-M. Chen, and Z. Yang, “Benchmark Functions for the CEC 2008 Special Session and Competition on Large Scale Global Optimization, ” *Nature Inspired Computation and Applications Laboratory*, *USTC, China*, 2007.

[22] R. Arora, R. Tulshyan, and K. Deb, “Parallelization of Binary and Real-Coded Genetic Algorithms on GPU using CUDA, ” in *2010 IEEE Congress onEvolutionary Computation (CEC)*, 2010, pp. 1–8.

[23] M. Oiso, T. Yasuda, K. Ohkura, and Y. Matumura, “Accelerating Steady-State Genetic Algorithms based on CUDA Architecture, ” in *2011 IEEE Congress onEvolutionary Computation (CEC)*, 2011, pp. 687–692.

[24] F. Stentiford, “An Evolutionary Programming Approach to the Simulation of Visual Attention, ” in *Proceedings of the 2001 Congress onEvolutionary Computation,* vol**.** 2, 2001, pp. 851– 858.

[25] Stefek, “Benchmarking of Heuristic Optimization Methods, ” in 14th *International SymposiumMECHATRONIKA*, 2011, pp. 68–71.

[26] R. Green, L. Wang, M. Alam, and R. A. Formato, “Central Force Optimization on a GPU: A Case Study in High Performance Metaheuristics using Multiple Topologies, ” in *2011 IEEE Congress onEvolutionary Computation (CEC)*, 2011, pp. 550–557.

[27] J. Jaros, “Multi-GPUIsland-based Genetic Algorithm for Solving the Knapsack Problem, ” in *2012 IEEE Congress onEvolutionary Computation (CEC)*, 2012, pp. 1–8.

[28] C. NVIDIA, “CUDACBest Practices Guide ver. 4.0, ” 2011.

[29] C. Toolkit, “4.0 CURANDGuide. nVidiaCorporation, version 12.3, January 2011.”

[30] J. K. Salmon, M. A. Moraes, R. O. Dror, and D. E. Shaw, “Parallel Random Numbers: As Easy as 1, 2, 3, ” in*2011 International Conference forHigh Performance Computing, Networking, Storage and Analysis (SC)*, 2011, pp. 1–12.

[31] J. Sanders and E. Kandrot, *CUDA by Example: An Introduction to General-Purpose GPU Programming*. Addison-Wesley Professional, 2010.

[32] P. Kromer, J. Platos, V. Snasel, and A. Abraham, “A Comparison of Many-Threaded Differential Evolution and Genetic Algorithms on CUDA, ” in *2011 Third World Congress onNature and Biologically Inspired Computing (NaBIC)*, 2011, pp. 509–514.

[33] S. Ali, T. D. Braun, H. J. Siegel, and A. A. Maciejewski, “HeterogeneousComputing, ” 2002.

[34] T. D. Braun, H. J. Siegel, N. Beck, L.L. B¨ol¨oni, M. Maheswaran, A. I. Reuther, J. P. Robertson, M. D. Theys, B. Yao, D. Hensgen *et al.*, “A Comparison of Eleven Static Heuristics for Mapping a Class of Independent Tasks onto Heterogeneous Distributed Computing Systems*, ” Journal of Parallel and Distributed computing,* vol. 61, no. 6, pp. 810–837, 2001.

[35] K. Price, R.M. Storn, and J.A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization.* Springer, 2006.

DDoS Attacks Impact Analysis on Web Service Using Emulation

Daljeet Kaur[1] and Monika Sachdeva[2]
[1]*Department of Computer Science & Engineering,*
SBS State Technical Campus, Ferozepur, Cantt–152001, India
E-mail: [1]Daljeetkaur617@gmail.com, [2]Monika.sal@rediffmail.com

***Abstract*—Banking, transportation, power, health, and defense are essential services being operated and these operations now days are being replaced by affordable and easily accessible Internet-based applications. It is all because of rapid growth and success of Internet in every sector. Unfortunately with it's the rapid growth, count of attacks has also increased incredibly fast. A Denial of Service (DoS) attack is a malicious effort to keep endorsed users of a website or web service from accessing it, or limiting their ability to do so. A Distributed Denial of Service (DDoS) attack is another type of DoS attack in which many computers are used to cripple a web page, website or web-based service. The services are severely degraded and hence lot of business loses are incurred due to these attacks. To objectively evaluate DDoS attack's impact, and the effectiveness of a potential defense, we need precise, quantitative and comprehensive DDoS impact metrics that are applicable to web services. To meet this requirement, the cyber-DEfense Technology Experimental Research (DETER) testbed has been developed. In this paper, we have created dumb-bell topology and generated background traffic as Web traffic. Different types of DDoS attacks are also launched along with Web traffic by using attack tools available in DETER testbed. Finally impact of DDoS attack on Web server is measured in terms of metrics such as throughput, percentage link utilization, and normal packet survival ratio (NPSR).**

Keywords: Terms-Internet, Distributed Denial of Service Attack, throughput, Percentage Link Utilization, Attack Traffic, Legitimate Traffic

I. INTRODUCTION

The main objective of Internet was providing an open and scalable network, which could offer easy, fast and inexpensive communication mechanisms, it was indeed very successful in accomplishing this particular goal. During Internet design, the functionality aspect was of much concern rather than security, which leads to several security issues that create a room for various attacks on the Internet. Internet security can be defined in terms of confidentiality, authentication, message integrity and non-repudiation out of which Availability is one of its main aspect. Attacks such as denial of service and its variant distributed denial of service attack target the availability of services on the Internet. Threat to the Internet availability is a big issue which is hindering the growth of online organizations those rely on having their websites 100% available to visitors, users and customers. A Denial of Service attack is an attempt by a person or a group of persons to decay an online service. This can have serious consequences, especially for companies like Amazon and eBay which rely on their online availability to do business. Recently there have been some large scale attacks targeting high profile internet sites [1, 2, 3, and 4]. Consequently, there are now a lot of efforts being made to come up with mechanisms to detect and mitigate such attacks. Even though the first denial of service attacks did not take place a long time ago (tools that automate setting up of an attack network and launching of attacks, started appearing in 1998), there is an abundance of denial of service attacks that have been used. Broadly speaking the attacks can be of three forms. a) Attacks exploiting some vulnerability or bug in the software implementation of a service to bring that down. b) Attacks that use up all the available resources at the target machine. c) Attacks that consume all the bandwidth available to the victim machine. The third type of attacks is called bandwidth attacks. A distributed framework becomes especially suited for such attacks as a reasonable amount of data directed from a number of hosts can generate a lot of traffic at and near the target machine, clogging all the routes to the victim.

II. DDOS ATTACK OVERVIEW

The normal functionality of the Internet servers is disabled during DDoS Attacks by exhausting resources. An attacker can create a huge volume of attack traffic and consume the bandwidth of the bottleneck link in the victim network to exhaust its resources. Due to the lack of application of security engineering in the development of operating systems and network protocols, hackers are provided with lot of insecure machines on internet. These insecure and unpatched machines are used by DDoS attackers as their army to launch attack [5]. An attacker or hacker gradually implants attack programs on these insecure machines. Depending upon sophistication in logic of implanted programs, these compromised machines are called Masters/Handlers or Zombies and are collectively called bots and the attacked network is called botnet in hacker's community. Hackers send control instructions to masters, which in turn communicate it to zombies for launching attack. The zombie machines under control of masters/handlers (running control mechanism), transmit attack packets as shown in Fig. 1, which converge at victim or its network to use up either its communication or computational resources [6].

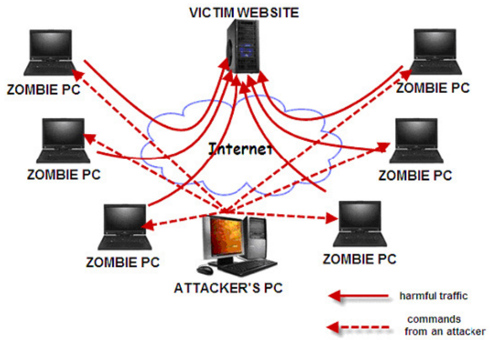


Fig. 1 DDoS Attack Architecture

Mirkovic *et al.* [6] and Peng *et al.* [7] have categorized DDOS attacks into two broad categories: flooding attacks and vulnerability attacks. Flooding DDoS attacks consume resources such as network bandwidth by overwhelming bottleneck link with a high volume of packets. Vulnerability attacks use the expected behavior of protocols such as TCP and HTTP as an advantage to the attacker. The computational resources of the server are tied up by seemingly legitimate requests of the attackers results in preventing the server from processing transactions or requests from authorized users.

Flooding DDoS is basically a resource overloading problem. The resource can be bandwidth, memory, CPU cycles, file descriptors and buffers etc. The attackers bombard the scarce resource(s) by utter flood of packets. In Figure 2, a flood of packets is shown, which congests the link between ISP's edge router and border router of victim domain [8].

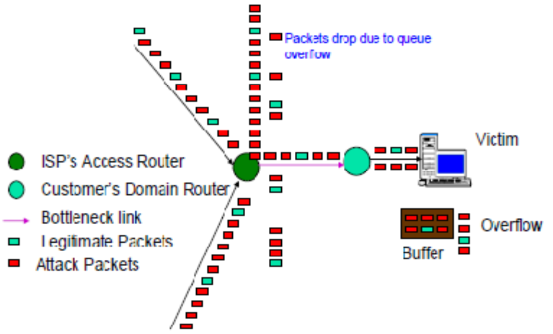


Fig. 2 Packets Drop During DDoS Attack

Attack packets keep arriving at user machine as per the distribution fixed by attacker, whereas legitimate clients cut short their packet sending rates as per flow control and congestion signals. A state comes when whole of bottleneck bandwidth is seized by attack packets. Thus, service is denied to legitimate users due to narrow bottleneck bandwidth. However, resources of connecting network are not a problem in case of commercial servers as these are hosted by the ISPs, quite close to their backbone network with high bandwidth access links. But server resources such as processing capacity, buffer limit etc., are put under stress by flood of seemingly legitimate requests generated by DDoS attack zombies [8]. Each request consume some CPU cycles. Once the total request rate is more than the service rate of server, as shown in Figure 2, the requests start getting buffered in the server, and after some time incoming requests are dropped due to buffer over run. The congestion and flow control signals [9], [10] force legitimate clients to decrease their rate of sending requests, whereas attack packets keep coming. Finally, a stage comes when only attack traffic is reaching at the server.

III. Related Work

The impact metrics of DDoS attack are closely related with measuring effectiveness of DDoS defense approaches. At present there are no benchmarks [11], [12] in terms of effective metrics for evaluating the impact and defense strategies of DDoS attack. Most of the existing strategies compare good-put under attack, without attack, and with defense [13]. Some of recent measurements [14] have also emphasized on response time. Evaluating the normal packets survival ratio proves to be the most important metrics as it clearly reflects accuracy of the defense and normal packet loss index [15], [16]. For measuring the impact of DDoS, Jelena *et al.* [17], [18] have used metric of percentage of failed transactions (transactions that do not follow QoS thresholds). They have defined an application specific threshold-based model for the relevant traffic measurements. When a measurement exceeds its threshold, it indicates poor services quality. But the absolute duration of threshold cannot be set since transaction duration depends on the volume of data being transferred and network load. Server timeout has been used as a metric in [19]. However collateral damage in terms of legitimate traffic drop is not indicated. Sardana *et al.* [20] have used good put, mean time between failure and average response time as performance metrics whereas Gupta *et al.* [21] have used two statistical metrics namely, Volume and Flow to detect DDoS attacks. As per [17] metrics such as good-put, bad-put, response time, number of active connections, ratio of average serve rate and request rate, and normal packet survival index [16] properly signal denial of service for two way applications such as HTTP, FTP and DNS, but not for traffic like media applications that is sensitive to one-way delay, packet and jitter.

IV. Experiment Setup

We used SEER (Security Experimentation Environment) GUI BETA6 environment [22] [23] to evaluate our metrics in experiments on the DETER testbed using. The test bed is located at the USC Information Sciences Institute and UC Berkeley, and allows security researchers to evaluate attacks and defenses in a controlled environment.

A. *Experimental Topology*

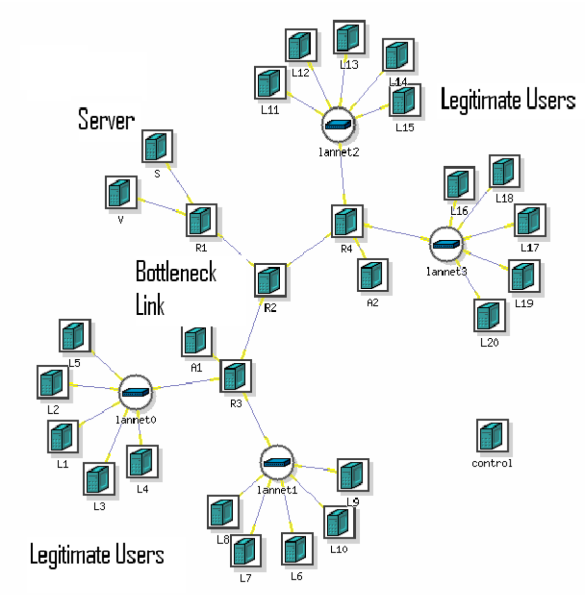


Fig. 3 Experimental Topology

Figure 3 shows the experimental topology and Figure 4[24] shows our experimental topology definition for Web applications in which R1, R2, R3 and R4 are routers, node S is server and L1-L20 are clients. They send legitimate requests to server S via router R1 and R2.

```
set ns [new Simulator]
source tb_compat.tcl
#Create the topology nodes
foreach node { V S R1 R2 R3 R4 L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 L14 L15 L16 L17 L18 L19 L20 A1 A2 control }
{
#Create new node
set $node [$ns node]
#Define the OS image
tb-set-node-os [set $node] FC4-STD
#Have SEER install itself and startup when the node is ready
tb-set-node-startcmd [set $node] "sudo python /share/seer/v160/experiment-setup.py Basic"
}
#Create the topology links
set linkRV [$ns duplex-link $V $R1 100Mb 3ms DropTail]
set linkRS [$ns duplex-link $S $R1 100Mb 3ms DropTail]
set linkRA1 [$ns duplex-link $A1 $R3 100Mb 3ms DropTail]
set linkRA2 [$ns duplex-link $A2 $R4 100Mb 3ms DropTail]
set linkRR3 [$ns duplex-link $R2 $R3 100Mb 3ms DropTail]
set linkRR4 [$ns duplex-link $R2 $R4 100Mb 3ms DropTail]
set linkRR2 [$ns duplex-link $R2 $R1 1.5Mb 0ms DropTail]
set lannet0 [$ns make-lan "$L1 $L2 $L3 $L4 $L5 $R3" 100Mb 0ms]
set lannet1 [$ns make-lan "$L6 $L7 $L8 $L9 $L10 $R3" 100Mb 0ms]
set lannet2 [$ns make-lan "$L11 $L12 $L13 $L14 $L15 $R4" 100Mb 0ms]
set lannet3 [$ns make-lan "$L16 $L17 $L18 $L19 $L20 $R4" 100Mb 0ms]
$ns rtproto Static
$ns run
```

Fig 4 Experimental Topology Definition

The bandwidth of all links is set to be 100 Mbps, and the bandwidth of bottleneck link (R1-R2) is 1.5 Mbps. Node A1 in topology acts as attacking node and it sends attack traffic to server S via router R1 and R2. The link between R1 and R2 is called bottleneck link. The purpose of attack node is to consume/congest the bandwidth of bottleneck link so that legitimate traffic could not get accessed by the server S.

We have generated a random network consist of Web clients, servers and attack source. In our emulated network, multiple legitimate clients connected with server and one attack source is used as DDoS flooding attacker. This emulates the real situation of DDoS flooding attack.

B. *Legitimate Traffic*

In our experiment, Web traffic is used where we have used 20 legitimate client nodes which send requests to the server S for 1-30 seconds and then 61-90 seconds with following thinking time. The configuration of said traffic parameters used to send legitimate traffic is demonstrated in Table I:

TABLE 1 EMULATION PARAMETERS USED IN EXPERIMENT

Parameter	Values
Client	L1-L20
Server	S
Attack Host	A1
Thinking Time	Minmax(0.01,0.1)
File Size	Minmax(512,1024)
Emulation Time	90sec
Bottleneck Bandwidth	1.5Mb
Access Bandwidth	100Mb
Legitimate Request Time	1-30 sec and 61-90 sec
Attack Time	31-60 sec
Attack Type	DDoS Packet Flooding
Server Delay	3ms
Access Link Delay	3ms
Backbone Link Delay	0ms

C. *Attack Traffic*

We have generated DDoS attack by using packet flooding attack. Node A1 launches attack towards S and thus consumes bandwidth of bottleneck in link R1-R2. UDP protocol is used for launching attacks. Further attack types flat, ramp-up, pulse and ramp-pulse are used in our experiment. Attack traffic from A1 starts at 31st second and stops at 60th second. Then we have analyzed impact of DDoS attacks on Web service. Table II shows attack parameters used in our emulation experiment. We have generated following flooding attack types:

Flat Attack: The high rate is achieved and maintained till the attack is stopped.

Ramp-up Attack: The high rate is achieved gradually within the rise time specified and is maintained until the attack is stopped.

Ramp-down Attack: The high rate is achieved gradually and after high time it falls to the low rate with in low time.

Pulse Attack: The attack oscillates between high rate and low rate. It remains at high rate for high time specified and then falls to low rate specified for the low tie specified and so on.

Ramp-pulse Attack: It is a mixture of Ramp-up, Rampdown and Pulse attack.

TABLE 2 ATTACK PARAMETERS USED IN EXPERIMENT

Attack Type	Flooding	Flooding	Flooding	Flooding
Attack Source	A1	A1	A1	A1
Attack Target	S	S	S	S
Protocol	UDP	UDP	UDP	UDP
Length Min	50	50	100	50
Length Max	100	50	150	50
Flood Type	Flat	Ramp-up	Pulse	Ramp-Pulse
High Rate	500	300	200	200
High Time	100	5000	5000	5000
Low Rate	300	100	50	50
Low Time	0	7000	4000	7000
Rise Shape	0	1.0	0	1.0
Rise Time	0	10000	0	10000
Fall Shape	0	0	0	1.0
Fall Time	0	0	0	10000
Sport Min	57	57	57	57
Sport Max	57	57	57	57
Dport Min	1000	1000	1000	1000
Dport Max	2000	2000	2000	2000
TCP Flags	SYN	SYN	SYN	SYN

V. RESULTS AND DISCUSSIONS

The effect of DDoS attacks on the performance of FTP service is analyzed below:

A. *Througput*

A backbone link is attacked to force the edge router at the ISP of victim end to drop most legitimate packets during a DDoS attack. In Figure 5 and Figure 6 we have concentrated on the throughput in terms of good-put and bad-put to get the measure of actual loss. So throughput is divided into good-put and bad-put respectively. Good-put is defined as no. of bits per second of legitimate traffic that are received at the server whereas bad-put gives no. of bits per second of attack traffic that are received at the server.

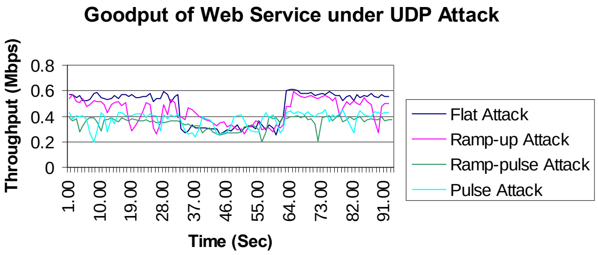


Fig. 5 Good-put of Web Traffic through Bottleneck Link During UDP Attack.

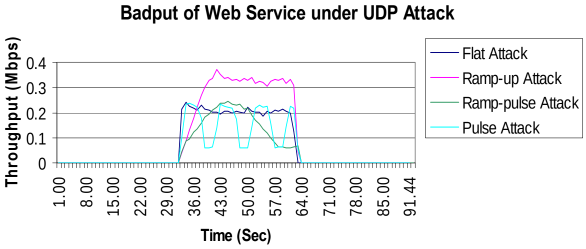


Fig 6 Bad-put of Web Traffic Through Bottleneck Link during UDP Attack

B. *Backbone Link Utilization*

Backbone Link utilization is defined as percentage of bandwidth that is carrying legitimate traffic. As shown in Figure7 Backbone Link Utilization is nearly 100% without attack. During attack, it drops more than 50%.

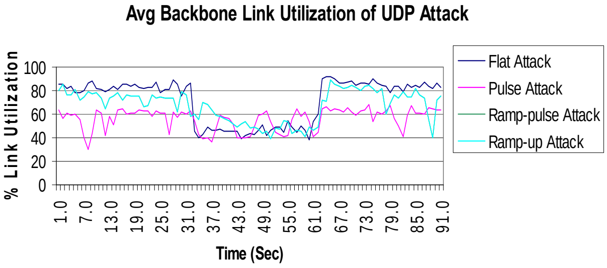


Fig. 7 Average Bottleneck Bandwidth Utilization in Web Service

C. *Normal Packet Survival Ratio (NPSR)*

NPSR is defined as ratio of good-put and bad-put. This is percentage of legitimate packets that can survive during attack. NPSR should be high. We can measure impact of attack as a percentage of legitimate packets delivered during the attack. If this percentage is high, service continues with little interruption. NPSR starts decreasing with increased rate of attack traffic and as bandwidth of the link is limited, so legitimate packets starts dropping. As shown in Figure 8100% legitimate packets are delivered without attack but during attacks, only 50% legitimate packets are delivered.

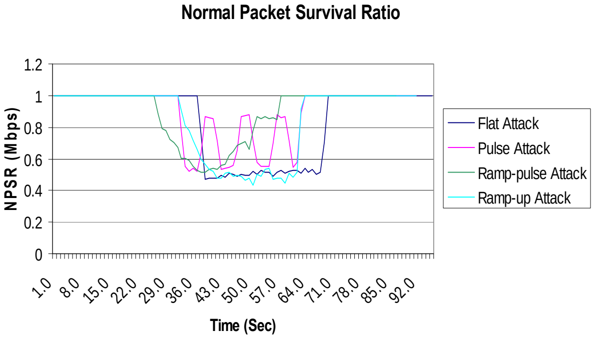


Fig. 8 Average Ratio of Legitimate Web Packets Survival during UDP Attack

ACKNOWLEDGMENT

We would like to express our gratitude to all those who gave us the possibility to complete this experimental work. We are extremely thankful to all the colleagues and faculty members for their constructive criticism and guidelines.

REFERENCES

[1] CNN. Cyber-attacks batter Web heavyweights, February2000.
[2] CNN.Immense. Network assault takes down Yahoo, February
[3] Netscape. Leading web sites under attack, February 2000 “Journal of Computer Science
[4] CERT coordination center. Denial of Service attacks
[5] J. Mirkovic. D-WARD: Source-End Defense AgainstDistributed Denial-of-service Attacks, Ph.D. Thesis,University of California, Los Angeles, 2003.
[6] J. Mirkovic and P. Reiher. “A Taxonomy of DDoSAttack and DDoS Defense Mechanisms”, ACM SIGCOMM Computer Communications Review,Volume 34, Issue 2, pp. 39-53, April, 2004.
[7] T. Peng, C. Leckie, and K. Ramamohanarao. “Survey of Network-Based Defense Mechanisms ountering the DoS and DDoS Problems”, ACM Computing Surveys, Vol. 39, No. 1, Article 3, April 2007.
[8] K. Kumar, R.C. Joshi, and K. Singh. “An Integrated Approach for Defending against distributed Denial-of-Service (DDoS) Attacks”, IRISS-2006, IIT Madras.
[9] M. Kisimoto. Studies on Congestion ControlMechanisms in the Internet–AIMD-based WindowFlow Control Mechanism and Active Queue Management Mechanism, Master Thesis, Osaka University, 2003.
[10] S. Floyd and K. Fall. “Router Mechanisms to Support End-to-End Congestion Control,” Lawrence Berkeley Laboratories Technical Report, 1997.
[11] J. Mirkovic and P. Reiher, A University of Delaware Subcontract to CLA.
[12] J. Mirkovic, E Arikan, S. Wei, R. Thomas, S. Fahmy, and P. Reiher. “Benchmarks for DDOS Defense Evaluation”, In Proceedings of Military Communications Conference (MILCOM), pp. 1-10, 2006.
[13] Y. You. “A defense framework for flooding based DDoS Attacks”, M.S. Thesis, Queen’s University, Canada.
[14] J. Mirkovic,P. Reiher,S. Fahmy,R. Thomas, A. Hussain, S. Schwab. “Measuring denial Of service”, 2nd ACM workshop on Quality of protection QoP, pp. 53–58, 2006.
[15] A. Hussian, S. Schwab, R. Thomas, S. Fahmy, and J. Mirkovic. “DDoS Experiment Methodology”, DETER Community Workshop, June 15-16, 2006.
[16] K. Kumar. Protection from Distributed Denial of Service (DDoS) Attacks in ISP Domain, Ph.D. Thesis, Indian Institute of Technology, Roorkee, India, 2007.
[17] J. Mirkovic, A. Hussain, B. Wilson, S. Fahmy, P. Reiher, R Thomas, W. M. Yao, S Schwab. “Towards user-centric metrics for denial-of-service measurement” in proceedings of the 2007 workshop on Experimental computer science, San Diego, California.
[18] J. Mirkovic, S. Fahmy, P. Reiher, R. Thomas, A. Hussain, S. Schwab,and C. Ko. “Measuring Impact of DoS Attacks”In Proceedings of the DETER Community Workshop on Cyberscurity,Experimentation, June 2006.
[19] C. Ko, A. Hussain, S. Schwab, R. Thomas, and B. Wilson. “Towards systematic IDS evaluation", in Proceedings of DETER Community Workshop, pp. 20-23, June 2006.
[20] A. Sardana and R.C. Joshi, “An Integrated Honeypot Framework for Proactive Detection, Characterization and Redirection of DDoS Attacks at ISP level,” International Journal Of Information Assurance and Security (JIAS), 3 (1), pp. 1-15, March 2008.
[21] B.B. Gupta, R. C. Joshi, and M. Misra, “An ISP Level Solution to Combat DDoS Attacks using Combined Statistical Based Approach,” Journal of Information Assurance and Security 3(2), 102-110, June 2008.
[22] M.Sachdeva, G.Singh, K.Kumar, K.Singh,”Journal of Information Assurance and Security 5 (2010)”, pp. 392-400. International Conference on Microelectronics, Communication and Renewable Energy (ICMiCR-2013).
[23] T. Benzel, R. Braden, D. Kim, C. Neuman, A. Joseph, K. Sklower, R. Ostrenga, and S. Schwab, “Experiences With DETER: A Testbed for Security Research.
[24] D. kaur, M. Sachdeva and K. Kumar,” Impact Analysis of DDoS Attacks on FTP Services” International Conference on Recent Trends in Information, Telecommunication and Computing, ISBN 978-94-91587-21-3,pp. 220-228, March 21, 2014.

Recent Flash Events: A Study

Avneet Dhingra[1] and Monika Sachdeva[2]
[1]Department of Computer Science and Engineering,
Punjab Technical University, Jalandhar, Punjab, India
[2]Department of Computer Science and Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]manavn@yahoo.com, [2]monika.sal@rediffmail.com

***Abstract*—With the evolving technology, the dependence on Internet, for business, communication and information exchange, has increased manifolds. Disruption of web services, even for small duration, leads to huge losses. The two major reasons for the disruption are DDoS attacks and Flash Events. Both cause the network to be overloaded, thus making the limited resources like-bandwidth, CPU, memory etc., unavailable to genuine users. Thus the need to find strategies to distinguish between the two arises. In this paper, we have explained flash events, their causes, effects, characteristics and also how they differ from DDoS attacks. The paper gives an explanation as to why should the server discriminate between DDoS attack and Flash Event. The Recent flash Events experienced by different websites have been presented so as to get the real world scenario of the same.**

Keywords: Flash Events, DDoS, Characteristics, Comparison of DDoS and Flash Events, Types of Flash Events, Slashdot Effect

I. INTRODUCTION

In the last decade technology has evolved manifolds, thereby changing the way of storing the information and accessing it. All the required communication takes place via Internet. Network links of the Internet act as conduit for transferring information. The society, at large, has started depending on web for business, research and related information flow. Disruption of services of web or malfunctioning of even a small part of network, for a small duration, degrades the performance and leads to huge losses [1].

The performance deterioration can occur due to two main reasons. First, it could be an intentional malicious attempt by attackers to disrupt the victim services. Such an attack is known as Distributed Denial of Service (DDoS) attack. DDoS are conducted using massive botnets which in turn use compromised servers known as Slaves or Zombies. These attacks overwhelm the network resources (CPU, Memory or network bandwidth) with requests, such that their services are rendered unavailable to the legitimate user.

Second, it could be a Flash Crowd which occurs when there is a sudden increase in volume of web traffic such that the response time of a website increases and in some cases also leads to the crash down of the affected website.

Both these situations arise due to variation in volume of the Internet traffic. Both of these lead to inconsistent behavior of the victim to the requests received.

This paper discusses *Flash Events* in detail and compares it with its counterpart-*DDoS attacks*. The major contribution of the paper is to provide an in-depth study of Flash events, its characteristics and its comparison with DDoS attacks.

Section II and III give the definition of flash events and their types. Section IV gives the effects of flash events. Section V describes the features of flash event. The need for discriminating flash events and DDoS attacks is highlighted in Section VI. Comparison of flash events and DDoS are given in section VII. In section VIII, world scenario of flash events is recounted. Section IX concludes the paper.

II. WHAT ARE FLASH EVENTS

The term Flash Event (aka Flash Crowd), for the internet, was inspired by Larry Niven's science fiction short story, "The Flight of the Horse", published in early 1970's. In this story, teleportation machine was invented which could take people back to the time in history when the major event occurred. However, author did not anticipate that huge crowd would teleport themselves to watch a certain event, and that it would lead to confusion and chaos at that particular place of an event.

In today's world of internet, the term is used to describe exponential rise in website traffic, when large number of users send the request for services simultaneously to the website which gives the details of an event. Such a surge leads to performance deterioration [2]. Events causing huge traffic could be some internationally acclaimed sports event like Olympics, Football World cup or release of new product by Apple or Microsoft. It can also occur in case of a natural disaster or a terrorist attack (example: 9/11 attack on America). Sometimes, a low efficiency server is linked to a very popular website like Slashdot or reddit, which may cause huge growth in traffic. Such a flash event is known as Slashdot effect [9].

Sachdeva *et al.* [3], describes flash crowds as "sharp and often overwhelming increase in number of users attempting to access a web site simultaneously in response to some event or announcement". Events which attract flash crowds can be referred as flash events. According to Bhatia *et al.* [1], it is used to describe a situation in which hundreds and thousands of valid users access a computing resource simultaneously. The computing resources could be CPU, network bandwidth or memory. Yu *et al.* [4], describes Flash crowds as "unexpected but legitimate, dramatic surges of access to a server." Wendell and Freedman [5] have explained FC mathematically. For an affected website, it is a time period over which request rate tends to increase exponentially.

If $r_{ti} > 2^{i}.r_{t0}$, $\forall i \in [0,k]$, the website is experiencing a flash crowd, where,

r_i =average per minute request rate over time $t_{i.}$

An event is said to be flash event only if the web server it affects needs to adjust the operation in order to remain available to users.

III. Types of Flash Events

Some flash events can be anticipated well before their occurrence and the sites can prepare themselves accordingly. These are known as *predictable flash events*. The world cup generally experiences a huge surge in the internet traffic. The related websites prepare themselves well in time to take care for such an increase in requests. However, despite number of preparations, Twitter.com faced an outage for 30 minutes during World Cup 2010. Other events like Olympics, online registration of a national level entrance test Gate/JEE etc., declaration of results of similar exam and launch of I-Phone 5 by Apple are a few of the predictable flash events.

The *unpredictable flash events* are the ones where the owners of the affected website are caught unawares. This generally happens in case of some breaking news like terrorist attack, Earthquake, tsunami or an epidemic (like swine flu). Such an event occurred on 11 September, 2001, when CNN experienced a sudden surge in traffic along with other major news websites.

Another type of flash event, known as S*lashdotting or Slashdot effect*, occurs when low performing websites suddenly become popular after being mentioned on a popular website like Slashdot.com. Announcing the publication on Redhat and LNXY caused only a slight increase in the request rate, and the actual flash crowd started when the article was linked by Linux Today and Slashdot. Interestingly, the form of the announcement influenced its resulting traffic surge. Linux Today published a complete article text, whose copy was hosted on the Linux Today server. Slashdot, in turn, published only a hyper link to the original article hosted on the origin server. As a consequence, announcing the article on Slashdot caused a distinctly larger traffic surge compared to publishing the article text on Linux Today. Another important observation is that the traffic surge was sudden, but not instantaneous. The request rate increased from about 30 requests per minute up to over 250 requests per minute within 15 minutes. This observation underlies the assumption that one can predict flash crowds by analysing the trends in request rate [9].

According to Chandra *et al* [15], flash events can also be classified according to load growth rate (time it requires to reach the peak from normal request rate), peak load (what is the maximum traffic it achieves) and duration (for how much time the high load of requests was experienced).

IV. Effects of Flash Events

Flash event generally occurs at an application level. Whenever flash event occurs, HTTP request rate increases suddenly. However, response rate to such requests decreases substantially. In extreme cases the web server may even crash.

Flash events exert heavy load, upto tens or hundred times more than the normal, on target web server, causing the server temporarily unreachable. Due to its overall unpredictability and relatively short duration, the traditional server side provisions lead to under utilization of resources [6].

V. Characteristics of Flash Events

The DDoS attacks and Flash events both lead to the disruption of the services to legitimate users. Thus, it is important to study their features so as to get the in depth knowledge. The overview of characteristics helps to develop a good intuition about what flash events are and how they come into existence [9]. This helps the server owners to proactively prepare themselves for such events.

There is a substantial increase in requests/web traffic up to hundred times more than average request observed daily. This surge in web traffic causes the performance to decline, connections to drop, and sometimes even crash the affected server [3].

This up rise in request rate is, however short lived. When the legitimate clients experience a low performance of website, they stop sending further requests. Gradually, the traffic surge returns to usual levels.

Most of the client requests are generated by the users who belong to same network or who have visited the page before. In other words, the requests are from users known to the server.

Also number of unique traffic clusters is quiet less as compared to source addresses. The requests received by web server follow a zipf-like distribution.

VI. Why Discriminate Flash Events from DDoS Attacks

Flash events and DDoS occur due to sudden and large surge in web traffic. They are quite similar in terms of network traffic phenomenon and both lead to disruption of services. However, the sources of requests received in case of DDoS are not legitimate (from the zombies or slaves), whereas in case of Flash events, they are from genuine users. Thus, it is important to differentiate them, or else, false alarms may be raised. It is a big challenge before the defenders as, if the data interpretation goes wrong, it may cause serious consequences [17]. The detectors may declare legitimate crowd of requests to be DDoS and vice-versa. Also, the techniques to be used for the mitigation of traffic are different for both the cases. In case of flash event, the websites need an internet-wide infrastructure support which is available publicly such as Content Distributed Network (CDN), server proxies and multilevel caches. Some part of the payload is shifted to CDN's or caches so that maximum number of requests can be responded to. In case the spike in traffic is due to DDoS, then different strategies come to play. Main aim in case of DDoS attack is not to respond to illegitimate users. This can be done in number of ways-using graphical puzzles, analysing user browsing dynamics, using honey pots and other methods that help server differentiate between the genuine users and illegitimate user request.

There is a motivation behind every DDoS attack. Motive could be anything from financial gains to political achievements or it could also be just to put forward the views of the attacker group. To obtain the maximum results, attackers use distributed and large number of botnets. They can even mimic flash events or take advantage of the event and send malicious data under the radar. Whenever a spike in internet traffic is detected by server, the first thing to be done is to look at certain parameters to make sure it is flash event. Genuine requests are sieved from the data received and malicious ones are ignored. That raises the need for discriminating flash event from DDoS attack.

As is clear from above discussion, in order to provide continuous service to legitimate users, it is important to study features of both flash events and DDoS and differentiate between the two. The major difference between them lies in their nature and origin such as their access intents and the distributions of their source IP address and the increased and decreased speeds of traffic between them [7, 10].

VII. Comparison of Flash Events and DDoS Attacks

DDoS and Flash Event are voluminous, bursty and unstable. They both cause high rise in network traffic and lead to disruption of services to legitimate users. Studying the differences between the two, help develop effective prediction and defense mechanism.

According to Jung *et al.*, flash events and DDoS have following differences. During Flash events, clients can be effectively aggregated into clusters. In fact, many have been registered in logs. In case of DDoS, the distribution of DoS attackers is geographically distributed in form of Zombies. Very few previously seen clusters are involved [8].

There is a decline in per client request rate during flash event but in case of DDoS there is no change in per client request rate during the surge. In case of flash event, the volume of traffic generated fluctuates and forms random zigzag wave as there is dynamic change in users, whereas the volume of DDoS attack remains stable throughout the attack [10].

Figure 1 and Fig. 2 consist of model graphs of Flash Events and DDOS Attacks showing its various features. Difference in the traffic pattern in case flash event and DDoS attack is clearly visible in the figures, thus, helping to understand their characteristics.

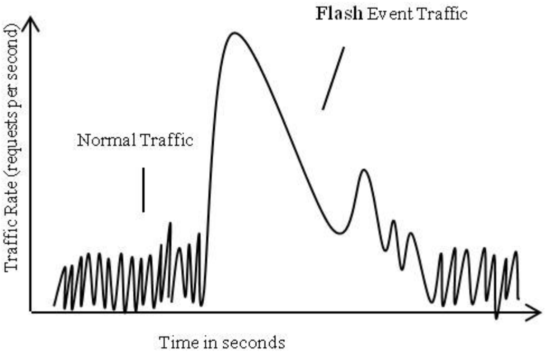


Fig 1 Model Graph for Flash Event

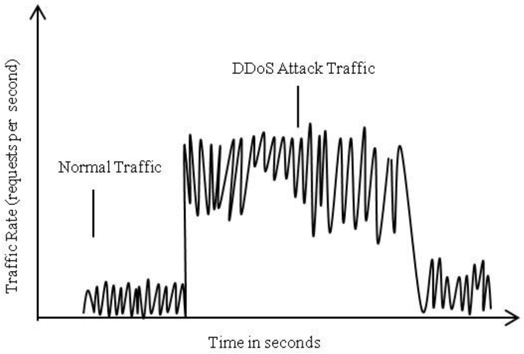


Fig. 2 Model Graph for DDoS Attack

Figure 1 shows that flash events grow rapidly and die out gradually. This is because the Event like any breaking news gets the requests suddenly. As soon as the user realizes the slow response rate, it stops accessing the affected server. After sometime, the Flash crowd declines. Also after certain time, the news has been known and accessed by all interested users. So, the news no longer attracts users, thus, decreasing the traffic.

Figure 2 shows the DDoS model graph depicting sudden rise and sudden fall of requests. It is so because DDoS attacks are conducted using botnets.

In short, the Flash events occur when there is breaking news or a world-wide event. In such a case, large numbers of users throughout the world, send requests to the web server for information. The sudden demand of information leads to outage or crash in the system. DDoS attacks are, however, well planned and programmed using the compromised systems known as zombies/ slaves. Therefore, the starting time and ending time are already defined.

Table I gives the comparison of DDoS attacks and Flash events.

TABLE 1 COMPARISON OF DDOS ATTACK AND FLASH EVENT

DDoS Attack	Flash Events
Network and server get congested and overloaded with the requests	Network and server get congested and overloaded with the requests
The traffic received is malicious and there is no need to respond.	The traffic received is genuine and need to be responded.
DDoS Attacks are always unpredictable. These occur as per the plans of an attacker using network of zombies	Flash Events can be predictable as well as unpredictable. These generally occur in case of a major world event like Olympics, Presidential elections, etc.
Request rate remains the same throughout the attack as the Zombies responsible for it, use automated tool in order to generate traffic.	Request rate per client decreases when compared to general state as the overload at servers result in drops and this forces request rate to drop at client side [3].
The requests received do not follow any particular pattern	The requests follow zipf-like distribution.

VIII. RECENT FLASH EVENTS

The flash events cause the server to get over whelmed with request. This sometimes causes the server's performance to decrease drastically and sometimes leads to crash of server. All this affects the clients using the web services. Frequent outages can lead to decrease in the number of users using the web site. Therefore, server owners use possible new methods to mitigate these events. The main aim of technical engineers, in such a scenario, is to restore the services at the shortest time possible.

The last decade has seen large number of flash events resulting in website outages. In this section, the recent Flash Events have been categorized according to the reason of traffic surge.

A. *Flash Events Due to Natural Disasters*

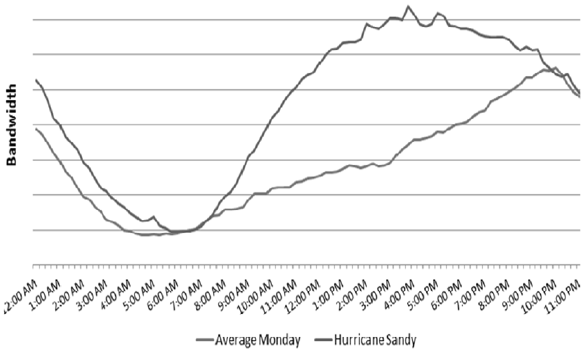


Fig. 3 Internet Traffic During Hurricane 'Sandy'. (*Source*: Sandvine)

In year 2012, super storm *Sandy* hit the eastern coast of The United States. The internet usage on 31 October increase by 114%. Netflix witnessed a traffic volume increase of 150%, while Skype witnessed a service usage increase of 122%, with a notable spike around 5pm. Fig. 3 below shows the East coast (USA) internet traffic for the day.

B. *Flash Events Due to Sports*

The 2010 World Cup, in South Africa, had the internet traffic exceed all the previous records. The leading social website, twitter, became the major victim. Normally it saw 750 tweets per second on an average day, but the traffic rose to approx. 2,940 tweets per second, whenever a goal occurred. These traffic spikes, overburdened the twitter's internal network capacity. It saw outages and maintenance downtime throughout the world cup [11].

Winter Olympics held in Sochi, Russia, saw a big rise in online traffic. The opening ceremony itself drove more than 1 Tbps of internet traffic.

C. *Flash Events Due to Launch of New Software/ Product*

According to Techcentral / Ireland's technology news resource, a unique breakdown occurred at Microsoft office, in June 2014 when Exchange Online and Lync Online, part of *Microsoft Office 360*, were unavailable for hours together. The previously unknown flaw had been detected in the directory partition due to which large number of customer could not access the email services. Even though connectivity was resumed, the resulting traffic surge overwhelmed the large number of network elements, thus leading to unavailability of Lync functionality for a little longer time.

On Sept. 18, 2013, Apple launched iOS7. Upon the release, Apple updates became almost 20% of total network traffic. Thousands of students at various universities in US (Ohio University, University of Texas, University of Arkansas), began to download it. This led to the surges as high as 5 times the normal traffic levels. Student newspapers also reported outages or slowdown of campus networks.

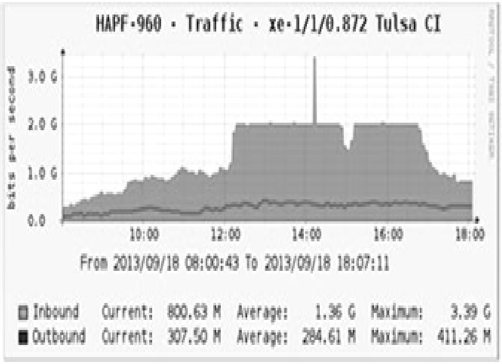


Fig. 4 Graph for Flash Event Occurred on Sept., 18, 2013 Due to Launch of iOS 7 Update on the University of Arkansas' Regional Arkansas Research and Education Optical Network Connection

Figure 4 gives an insight into the internet traffic spike caused by iOS 7's update launch. IOS 7 downloads caused the web traffic on Arkansas' regional Arkansas Research and Education Optical network connection to rise from 1.4 Gb/s to 6 Gb/s.

D. *Flash Events Due to Celebrities*

The websites of celebrities also sometimes get affected by Flash Events. In August 2013, the unusual trigger led to all the previous records of the Twitter's tweets-per-second to be destroyed. It was the broadcast of anime *master Hayao Miyazaki's* most famous movie "Castle in the Sky". Hundreds and hundreds of Japanese fans of the movie tweeted a magic-word used in the classic anime (short for animation), all at once. The word typed was "balse". It is spoken during movie's climax scene. The flood of tweets peaked at 143,199 tweets-per-second. The other websites like Amazon, Play station, KFC and Nissan, including "balse" button, experienced the failure as soon as the button was pressed [12]. Fig. 5 shows the traffic graph of the day (courtesy: blog.twitter.com). The spike shown is about 25 times the normal traffic.

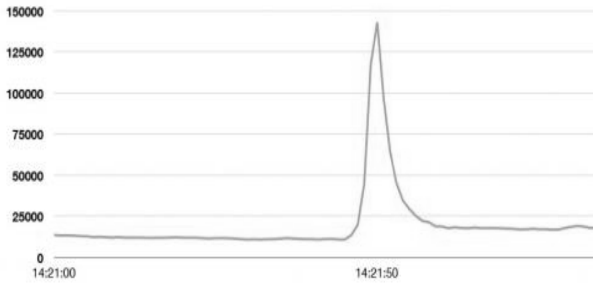


Fig. 5 Spike in Twitter Traffic Due to Sudden Typing of Word 'Balse' by Japanese People. (*Source*: blog.twitter.com)

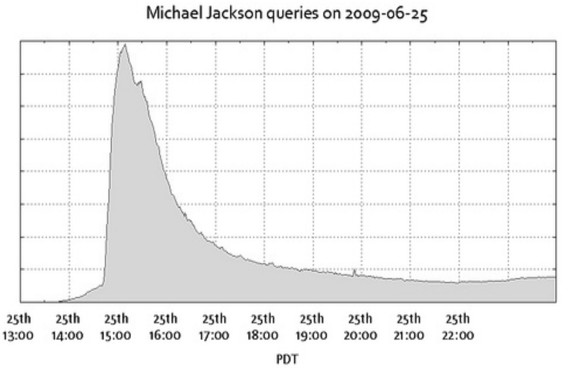


Fig. 6 Internet Spike at Google Search Engine Caused After Michael Jackson's Death. (*Source*: searchengineland.com)

When the *Michael Jackson's* news site TMZ first broke the news of MJ's death in June 2009, people from all over the world accessed the website to confirm it. Due to huge amount of requests at TMZ, it began to experience erratic outages. Other news sites like ABC, CBS, LA Times, AOL, CNN money, also became unavailable (down to nearly 10% availability). Activity at twitter peaked with 25% of all tweets happening at the time the reports confirmed that Jackson had died. In UK, twitter had the busiest day ever. Fig 6 shows the spike in traffic spike at Google when there was an outpour of searches related to Jackson. As we can see the traffic was too much that initially Google thought that it was under an attack.

According to data from Akamai, Internet traffic rose by 24% globally after the breaking news of *Osama Bin Laden's* Death. This high wave of web traffic, crippled CNN's news site. Other major news sites were also slowed down. Some of the requests were rejected also.

Salman Khan, the god-father of Bollywood, sent out a few tweets concerning his unemployed fans. The tweet said that he can talk to his friends to provide employment to his unemployed fans. Within a few minutes of tweeting, the stars Facebook page crashed due to heavy traffic from job-seekers.

E. *Flash Events Due to Other Reasons*

BestBuy, one of the USA's biggest electronic retailers, left its customers waiting for nearly an hour to complete the checkout process on November, 2012. It was due to huge surge in traffic to its site by the holiday shoppers. Users were greeted with the message: "we were expecting snow but we got a blizzard. Our site is incredibly busy! Please be patient while we shovel you a path" [13].

After the *London attacks* in July 2008, there was an up rise in the internet traffic leading to outages and degrading performances. At peak time, the BBC News website served up to 1.7 gigabits of data every second, with 40,000 page requests per second. The most affected news sites were BBC News, Netcraft, Sky News and MSN which were not available for short time.

IX. CONCLUSIONS

Studying about Flash events, help us get to know the features of flash events and features that differentiate it from DDoS attacks. This study helps to design the defense system which could predict the occurrence of spike in traffic at the web server so that pro-active measures can be taken to mitigate the huge traffic load on the server. The attacker may send malicious data to the server mixing it up with the flash event payload.

Studying the features of flash event helps in having a good insight about what exactly flash events are. Also it helps to know about their existence. Even when the web server owners expect traffic bursts, problems do occur due to non preparedness on the part of the owners or request per second surpassing the extra ordinary numbers. This causes losses due to outages or low performance. Technologies have to be at place in order to avoid outages. If it's a flash event then load balancing is required in order to maintain the response time, else if it is a DDoS attack, then the server need not respond to malicious requests but at the same time needs to take care of genuine users also.

The overall scenario, discussed in this paper, leads to the conclusion that traffic of both DDoS attack and flash event are unstable, voluminous, and occur in short bursts. Even the impact of both is similar. They both lead to complete or partial failure of the services provided by the affected server. Web server is required to identify and serve as many genuine requests as it can respond back to. Thus the need arises to discriminate the genuine users from illegitimate ones. It becomes necessary to see if the request is from some Zombie or slave, or it is the genuine user demanding the information. On the face of it, both requests seem to be coming from the authentic source. To distinguish them we need to learn about their characteristics and features that make them different. Thus, developing the technique to discriminate the data is a challenging job and needs a lot of in-depth study of the related information.

REFERENCES

[1] S.Bhatia, G. Mohay, A.Tickle, E.Ahmed, "Parametric differences between a real-world denial-of-service attack and a flash event," IEEE Computer Society, Sixth International Conference on Availability, Reliability and Security, 2011.

[2] Definition: Flash Crowd, available at, http:// www.catb.org/ jargon/ html/ F/ flash-crowd.html.

[3] M. Sachdeva, Thesis: "A distributed approach for defending the service against DDoS attacks," PhD. Thesis, Sept. 2012.

[4] S. Yu, W. Zhou, W. Jia, S.Guo, Y.Xiang, F.Tang, "Discriminating ddos attacks from flash crowds using flow correlation coefficient," IEEE Transactions on parallel and distributed systems, vol.23, No.6, pp 1073-1080, June 2012.

[5] P. Wendell, M. J. Freedman; "Going Viral" Flash Crowds in an open CDN", IMC'11, Nov.2-4, 2011; Berlin Germany. cs.princeton.edu/~mfreed/docs/flash-imc11.pdf

[6] C.PAN, M. Atajanov, M.B. Hossain, T. Shimokawa, N. Yashida, "FCAN: Flash Crowds alleviation network using adaptive P2P overlay of cache proxies," IEICE Trans. Communication, Vol.E 89-B, No.4, pp-1119, April 2006.

[7] Li. Ke, Z. Wanlei, L. ping; H. Jing, l. Jianwen, " Distinguishing DDoS attacks from Flash Crowds using probability metrics," NSS 2009: Proceedings of 3rd International Conference on Network and System Security, IEEE, pp 9-17

[8] J. Jung, B. Krishnamurthy, M. Rabinovich, "Flash Crowds and Denial of Service attacks: characterization and implications for CDNs and web sites," available at http:// www2. research.att.com/ ~bala/ papers/www02-fc.html.

[9] H.Izycka, "Flash Crowd prediction", Vrije Universiteit Amsterdam, Master's thesis, available at http:// www.globule.org/ publi/ FCP_master2006.pdf.

[10] K.M. Prasad, A.R.M. reddy, K.V. Rao, "Discriminating DDoS attack traffic from Flash Crowds on internet threat monitors (ITM) using entropy variations", AJC & ICT, IEEE, vol.6 6 No.2, June 2013.

[11] R. Miller, "Record world cup traffic slams twitter", available at datacenterknowledge.com, phys.org, June 2010

[12] T. Kontzer, "Twitter spike highlights need to plan for traffic surge", available at network computing.com, Aug., 2013.

[13] G. McQuaid, "How the cloud can help with seasonal traffic spikes", available at nexusbg.com, Dec., 2013.

[14] International Business, P. Goswami, "Salman Khan offers Jobs to Unemployed fans online; Resultant traffic crashes his Facebook Page", available at Ibtimes.co.in, June, 2014.

[15] A.Chandra, P. Shenoy, "Effectiveness of dynamic resource allocation for handling internet flash crowds," available at http://lass.cs.umass.edu/papers/pdf/TR03-37.pdf.

[16] S. Kandula, D.Katabi, M.Jacob, A. Berger, " Botz-4-Sale: surviving organised DDoS attacks that mimic flash crowds," Proceedings Second Symposium Networked Systems Design and Implementation(NSDI '05), 2005.

[17] S. Yu, T. Thapngam, J. Liu, S. Wei, W. Zhou, "Discriminating ddos flows from flash crowds using information distance," Proceedings of the 3rd International Conference on Network and System Security(NSS '09), pp 351-356.

[18] R. Saravanan, S. Shanmuganathan and Y. Palanichamy, "Behavior based detection of application layer distributed denial of service attacks during flash events."

[19] M. Sachdeva, K. Kumar, "A traffic cluster entropy based approach to distinguish DDoS attacks from flash events," Hindawi publishing corporation, ISRN Communication and networking, vol. 2014.

Survey of Stability Based Routing Protocols in Mobile Ad-hoc Networks

Mandeep Kaur Gulati[1] and Krishan Kumar[2]
[1]Punjab Technical University, PTU, Kapurthala, Punjab, India
[2]Department of Computer Science and Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]gulati_mandeep@rediffmail.com, [2]k.salujasbs@gmail.com

***Abstract*—A Mobile Ad hoc Network (MANET) is an autonomous collection of mobile nodes forming a dynamic network and communicating over wireless links. With the rising popularity of MANETs and demand of users, Quality of Service (QoS) has become major issue to be discussed. One of the most important criteria determining the assurance of QoS support in such networks is link stability. Due to the mobility of the nodes, link failures occur frequently and the route involving those links would no longer work. Stability therefore is an important element to be considered in the design of routing protocols. Stable paths, also called the long-lived paths, can thus be discovered and used to reduce the overhead resulted from route maintenance in ad hoc networks. A number of stability based routing protocols have been proposed in the literature. This paper presents the overview of the different approaches used to find the stable paths and a survey of some of the stability based routing protocols along with their strengths and weaknesses. Finally, a comparative study of all routing protocols is provided.**

Keywords: Mobile Ad hoc Network (MANET), Routing Protocol, Link Stability, Signal Strength, Residual Lifetime

I. INTRODUCTION

A Mobile Ad-hoc Network (MANET) is a collection of mobile nodes that form a dynamic infrastructure-less communication network wherever it is required. The nodes in the network not only act as hosts but also as routers that discover and maintain routes to other nodes in the network. Mobile Ad Hoc Networks (MANETs) are becoming the crucial medium of present day communication owing to their self-configuring, easily deployable and infrastructure-less nature. These networks are particularly suitable for emergency situations like warfare, floods and other disasters where infrastructure networks are impossible to operate. Since mobile nodes move in various directions causing existing links to break and the establishment of new routes, routing in such networks is a challenging task. Routing protocols used in these dynamic networks should be designed in such a way that they can adapt fast and efficiently to unexpected changes in network layout. Many routing protocols have been developed for mobile ad hoc networks such as Ad hoc On-demand Distance Vector Routing Protocol (AODV) [1] Destination–Sequenced Distance Vector (DSDV) protocol [2], Wireless Routing Protocol (WRP) [3], Temporally-Ordered Routing Algorithms (TORA) [4], Dynamic Source Routing Protocol (DSR) [5], Associativity Based Routing Protocol (ABR) [6], and Zone Routing Protocol (ZRP) [7], etc. These protocols tend to establish a path with least number of hops. Also, all these routing solutions only deal with the best-effort data traffic.

Currently a lot of applications have been developed for wireless networks, their practical implementation and use in the real world has been limited so far. Many of these applications such as real-time audio and video are sensitive to the Quality of Service (QoS). Hence focus has been shifted from best-effort service to the provision of better defined QoS in ad hoc networks. The most commonly employed QoS metrics [8] are link stability, link reliability, end-to-end delay, node buffer space, delay jitter, packet loss ratio etc. The parameter ‘link stability’ i.e. the predicted lifetime of a link is the most important criteria determining the assurance of QoS support. Node movements cause link breakages in MANETs. Thus instead of selecting weak links which may break soon and introduce maintenance overhead one can find path involving stable links i.e. having longer predicted lifetime. Stability or lifetime of a path is determined by the number of links that compose the path and the stability of each link in the path. Many stability based routing protocols have been proposed in the literature that enhance network stability. The primary goal of most stability based routing protocols is to find and select the paths that will last longer. These protocols reduce routing overhead and improve QoS performance as compared to the shortest path routing protocols. A few of the routing protocols along with their strengths and weaknesses have been discussed in the paper. Finally, a comparison of the routing protocols has been done so as to explore the future areas of work.

The rest of the paper is organized as follows. Section II presents overview of different approaches used to find the stable paths. In Section III, survey of some of the stability based routing protocols has been presented. In Section IV, a comparison of the routing protocols has been done. Finally, Section V concludes the paper and gives possible future directions in this research field.

II. OVERVIEW OF THE DIFFERENT APPROACHES TO FIND STABLE PATH

In a mobile environment, because of the mobility of mobile nodes in MANETs, the shortest path is not necessarily the best path. If the stability of a routing path is not considered, then wireless links may be easily broken. Link stability indicates how stable the link is and how long it can support communications between two nodes. Stability of links can be estimated using many parameters like-Signal Strength [15,17], hello packets or pilot signals [6, 11, 13], relative speed between two nodes forming the link (by Global Positioning System) [10, 12], Residual Lifetime [6, 10, 11, 12]. These different approaches or techniques have been discussed below:

The Global Positioning System (GPS) [9] is a very popular technique used to detect the exact position of the mobile nodes. Each node can calculate its position and a protocol is applied, which disseminates or requests the position for other nodes. There are routing protocols [10, 12] using the information obtained from the GPS. Many researches assume that GPS can be simply utilized in an open area environment. However, in the urban area, buildings, walls, and trees etc. may be there to form shields against the GPS signals. Moreover the use of GPS is greatly limited by shadow effect and multipath fading, and thus the GPS-aided routing protocols do not work well in such environment.

The Received Signal Strength (RSS) [15, 17] can be also taken as an index of link lifetime. In signal strength based approach, the receiving node captures the control packet and forwards it if and only if the packet signal strength is above a certain threshold otherwise discards the packet. This approach is easy and most efficient because it utilizes signal strength values from MAC layer to compute link stability. Since pilot signals are not exchanged periodically to compute link stability, it uses less control overhead. The received signal strength can be an accurate index to measure the link lifetime in the open area. However, in the urban area, a link may fail abruptly because of shadow effect, and the prediction error may thus increase.

In Hello Packet or Pilot Signal [6, 11, 13] based technique, each node periodically broadcasts a one-hop and ack-free hello packet to identify itself. By continuously receiving the hello packets, a node can verify the existence of its neighbours. Number of hello packet received by neighbour is used to determine the lifetime of the link. If a neighbouring node moves out of the radio coverage, the receiving of hello packets would terminate and therefore the node can recognize that a link has failed. When a link fails, the corresponding two nodes of the link record this link failure event on their lifetime records along with the lifetime of the failure link. The idea of this scheme is to establish routes over stationary nodes as possible to prevent frequent route failures. The link lifetime of all links to a stationary node tend to be longer and, in the contrast, the link lifetime of all the links to a moving node tend to be shorter. When a moving node becomes stationary, longer-lived links are emerging more likely. Thus if a link whose age is greater than the lifetime of these links being stored in the link lifetime record, then that link is considered to be reliable and stable.

Residual lifetime [6, 10, 11, 12] is also an efficient approach to find the stable route. In this approach, lifetime of the link is measured with the help of recent hello messages or relative velocity and direction of mobile nodes. This link lifetime is used to construct stable path. The residual lifetime of the path is the minimum link expiration time of node on that path and the path which has maximum link expiration time is selected as primary path. For n paths (π1, π2, ... πn) from source to destination, lifetime of a path is bounded by the lifetime of all the nodes along the path. When a node dies along a path we can say that the path does not exist any longer.

III. Survey of Stability Based Routing Protocols

Many of the stability based routing protocols have been proposed in the literature. A few of these routing protocols having different approaches for finding stable paths are surveyed in this section. For each protocol, the functionality and main features are described briefly.

A. *Associativity Based Routing (ABR)*

C-K. Toh [6] proposed Associativity Based Routing (ABR) protocol which is probably the first protocol in the class of stability based routing protocols for MANETs. It uses periodically sent pilot signals to determine the link stability. In this protocol, a metric called *associativity* is defined to determine the link stability. The protocol is based on the idea that nodes which have been stationary for a threshold period are less likely to move away. It assumes that after the threshold period, nodes move with similar speeds in similar direction and and thus tend to stay together for a longer period of time. The ABR protocol consists of three phases, namely route discovery phase, route re-construction phase and route deletion phase. Initially when a source node desires a route, the route discovery phase is invoked. The route discovery phase consists of a broadcast query which is broadcasted by the source. The intermediate node appends its address/identifier at the intermediate node ID field of the query packet and broadcasts it to its neighbours (if it has any). The associativity ticks with its neighbours will also be appended, along with its relaying load, link propagation delay and the hop count. The destination, at an appropriate time after receiving the first broadcast packet, knows all the possible routes and their qualities. It can then select the best route and send a REPLY packet back to the source, via the route selected. However, if the overall degree of association stability of two or more routes are same, then the route with the minimum hops will be chosen. If multiple routes have the same minimum-hop count, then one of the routes is arbitrarily selected. When the link of an established route changes due to source, destination, intermediate nodes or mobile hosts (MH's) migration, the route re-construction phase is invoked. When source no longer desires the route, the route deletion phase is initiated.

Simulation results show that the property of having long-lived routes enhances the communication throughput considerably and the capability of the routing protocol to quickly locate an alternative shorter route enhances the response time to link changes. One of the problems with ABR is the choice of the threshold value. This value may vary depending on the mobility patterns.

B. Flow Oriented Routing Protocol (FORP)

W. Su *et al.* [10] suggested an approach based on the availability of GPS measurements. The Flow Oriented Routing Protocol (FORP) follows an approach of calculating a link's residual lifetime from a mobile's own speed and the speed and distance of the connected party. However, this method strongly depends on the assumption of a free space propagation model and on having GPS equipment to estimate the expiration time of the link between two adjacent mobile nodes. When the sender has a flow to send, it constructs a route to the destination on demand and injects the flow. The destination predicts the change in topology ahead of time and determines on the route, the Route Expiration Time (RET) can also be predicted. Based on this prediction, routes are reconstructed before they expire. Simulation results indicate that with mobility prediction enhancements, more data packets were delivered to destinations while the control packets were utilized more efficiently. Since GPS may not work properly in certain situations (e.g., indoor, fading, etc.), the link expiration time for a particular link may not always be accurately predicted.

C. Stability and Hop-Count Based Approach for Route Computation (SHARC)

K.N. Sridhar and M. C. Chan [11] propose Stability and Hop-Count based Approach for Route Computation (SHARC) in MANET that considers both the hop-count and stability metrics. DSR (which is hop-count based) is used as the basic routing protocol and the residual link lifetime is calculated using a simple histogram based estimator. The protocol finds the most stable route among the set of shortest hop routes. In order to distribute stability information, the route-request packet of DSR is changed to carry residual lifetime information. Every node stores the link duration values of its neighbours. By collecting this information and aggregating them into bins of 10s, each node maintains an estimate of the residual lifetime distribution using the samples collected so far. When the intermediate node receives the route request packet, it includes the residual lifetime value in the packet. The path structure is changed by associating every path with an additional stability value. This stability value of the path is the sum of all the residual lifetime divided by the length of the path. The cache structure is also enhanced to maintain the stability metric along with the addresses of intermediate nodes. The route selection mechanism is incorporated in all the nodes so as to be compatible with DSR routing mechanism.

Simulation results show that it performs better than purely stability based and purely hop count based algorithms in terms of throughput of long-lived flows and response time of short data transfers.

D. Stable, Weight-based, On-demand Routing Protocol (SWORP)

N-C. Wang *et al.* [12] propose a stable, weight-based, on-demand routing protocol (SWORP). The protocol uses the weight-based route strategy to select a stable route in order to enhance system performance. The weight of a route is decided by three factors: Route Expiration Time (RET), Error Count (EC) and Hop Count (HC) where RET is the minimum link expiration time (LET) for a feasible path where LET represents the duration of time for a packet to travel between two nodes, EC captures the number of link failures caused by a mobile node and HC is the number of hops for a feasible path. All the nodes are assumed to have their clocks synchronized using the Global Positioning System (GPS) clock, so that two adjacent nodes may predict the RET. Route discovery usually first finds multiple routes from the source node to the destination node with the different weight values. Then the destination selects the path with the largest weight value for routing. The simulation results show that the protocol selects a stable routing path and reduces the routing overhead and packet loss. While the proposed scheme may fight against link breaks due to mobility, but it does not consider link breaks due to the draining node energy that must also be accounted for when computing weights for stable routing.

E. Stable and Delay Constraints Routing (SDCR)

P. Yang and B. Huang [13] proposed another Stable and Delay Constraints Routing (SDCR) protocol which extends the DSR protocol and adopts source routing mechanism. In the route discovery phase, the protocol finds paths that meets delay requirement with great link stability factor. In the route maintenance phase, it effectively keeps monitoring the network topology changes through delay prediction and performed rerouting in time. The SDCR includes two major phase namely routing discovery and routing maintenance. In the routing discovery process the SDCR find feasible paths between source and destination node while in the routing maintenance phase SDCR monitors and predicts the future information about availability of link. Link stability factor and delay constraints are taken into consideration in their route discovery and maintenance phases. In the SDCR, the RREQ of original DSR is extended and added to new fields namely delay constraint, time stamp and link stability factor coupled with the location and velocity of nodes. In the routing cache, link stability factor and delay constraint are added. In the route maintenance phase, SDCR effectively keeps monitoring network topology changes by delay prediction and performs rerouting before the paths become unavailable. The SDCR significantly improves routing performance with these route discovery and maintenance mechanisms operating together and it also guarantees QoS request.

The performance of SDCR was compared with the original DSR and DQR [14] and the results show that SDCR outperforms than other two protocols. It reduces the packet losses and guarantees the reliable and rapid transmission. Its advantage is remarkable in high mobility. However, the extra computation for link stability factor in SDCR causes the slightly higher delay.

F. Route Stability based QoS Routing (RSQR)

Sarma and Nandi [15] proposed an on-demand AODV based Route Stability based QoS Routing (RSQR) protocol in MANETs. The protocol uses route stability along with throughput and delay. The routing algorithm forwards the route request through all the feasible paths from source to destination avoiding very weak links during its forwarding process. To compute a QoS route to a destination *D*, the source *S* generates a QoS Route Request (QRREQ) packet with values for B_{min} and D_{max} from the application's requirements. An intermediate node *i*, after receiving a QRREQ packet, checks the signal strength of QRREQ and simply drops the packet if its strength is very poor (less than a threshold).Otherwise, node *i* performs the delay and throughput admission control. If the QRREQ passes both delay and throughput admission control, node *i* makes a temporary reverse route entry in *RT* (*Routing Table*). After the processing, some fields in QRREQ are modified such that the modified values contain the route stability and end-to-end delay of the explored route up to the current node. When the destination node receives the first route request, it waits for a fixed small time interval, called Route Reply Latency (RRL), for more route request packets to arrive. The destination would then select, among all feasible paths, the one with the highest route stability value to reply to the source. Therefore, the use of route stability during route discovery yields the route that last longer and consequently increases the throughput while reducing the delay and routing overhead.

The performance of the protocol was compared with AQOR [16] under different mobility and network load conditions and the results show that the RSQR protocol performs better than AQOR in terms of packet delivery ratio, routing overhead, end-to-end delay especially in high mobility conditions with marginal decrease in traffic admission ratio.

The drawback of the protocol is that it does not consider the issues like detection of potential link breaks or QoS violations before actual link breaks or QoS violation takes place. This results in performance degradation as the mobility of the nodes increases.

G. Routing Based on Multiple Constraints

D.S. Thenmozhi and M. Rajaram [17] presented multi constraint based routing technique to incorporate Quality of Service based applications in MANETs. AODV routing protocol is extended to perform path finding that meets the application stipulated bandwidth requirement and link stability metrics. During the route discovery process, the source broadcasts Route Request (RREQ) packet. It includes application's channel bandwidth requirement (BWflow) computed by the source, link stability indicator (Pr-fail, Tr) pair where Pr-fail represents the expected route break probability and Tr represents the expected time duration of the flow. Another field Pa is also evaluated and added which represents the accumulated survival probability of all the selected links from the source node to the current node. Then the node rebroadcasts the route request. Recording the sequence of hops in RREQ packet enables to determine the lower bound of the contention count of the complete route and also it can be used to eliminate circular routes.

When the intended destination receives a route request, it receives the full route and sends a route reply (RREP) back to the source along the same route. The destination may get different routes. The destination gives the preference to the route having all links possessing positive indication for the link stability. Simulation results prove that this approach of routing algorithm improves QoS performance in a significant way.

IV. Comparison of Stability Based Routing Protocols

The comparison of the above discussed routing protocols is shown in Table 1 below.

TABLE 1 COMPARISON OF STABILITY BASED ROUTING PROTOCOLS

Protocol	Base Protocol	Approach/Metric used to Find Stable Routes	Stability Parameter	Disadvantage	Mobility Support
Associativity Based Routing (ABR) [6]	DSR	Hello packet, Residual Lifetime	Association of neighbouring nodes	Assume that older links are more stable which is not always correct. Choice of threshold value is difficult as it varies depending on the mobility pattern.	Moderate
Flow Oriented Routing Protocol (FORP) [10]	_	GPS, Residual lifetime	Link expiration time calculated with the help of free space propagation model and GPS.	Strongly depends on the assumption of a free space propagation model and on having GPS equipment mobile node.	Moderate
Stability and Hop-Count based Approach for Route Computation (SHARC) [11]	DSR	Hello packet, Residual lifetime	Hop count, stability of a path calculated using a simple histogram based estimator	Path stability depends on average value of residual lifetime which is not efficient	Performs well in both low and high mobility
Stable, Weight-Based, On-demand Routing Protocol (SWORP) [12]		GPS, Residual lifetime	Weight function which includes link expiration time, error count and hop count	Depends on GPS which is not efficient in MANETs due to limited resources.	Moderate
Stable and Delay Constraints Routing (SDCR) [13]	DSR	Hello packet	Link stability factor with delay constraint	Extra overhead in DSR RREQ field	High mobility
Route Stability based QoS Routing (RSQR) [15]	AODV	Signal strength based	Signal strength	Complex calculation at each node	Both low and high mobility
Routing Strategy based on multiple constraints [17]	AODV	Signal strength based	Time count of the neighbouring nodes	Extra control overhead	Both low and high mobility

V. CONCLUSION

In this paper, the basic approaches and a brief description of a few of the stability based routing protocols in MANETs has been presented. The protocols are selected in such a way so as to highlight the different approaches to stable path routing in MANETs, while simultaneously covering most of the important advances in the field. A comparison of all the routing protocols has been provided and the strengths and drawbacks of these protocols have also been summarized so as to explore the future areas of research. However, routing protocols that are based only on link stability have either been shown to exhibit little improvement over hop-count based algorithm or the improvement comes when link lifetime can be accurately predicted. A crucial issue with stability based routing protocols is that much longer routes can be obtained as compared to hop-count based routing. Thus these protocols need to be further extended in the areas of multipath routing, load balancing, resource reservation, energy efficiency, security and cross layer design to improve their performance

REFERENCES

[1] C.E. Perkins, E.M. Royer and S.R. Das, “Ad hoc on-demand distance vector(AODV) routing protocol,” IETF Draft, RFC 3561, February 2003.

[2] C. Perkins and P. Bhagwat, “Highly dynamic destination-sequenced distance vector routing (DSDV) for mobile computers,” in *Proceedings of ACM SIGCOMM'94*, pp. 234-244, 1994.

[3] S. Murthy and J. J. Garcia-Luna-Aceves, “An efficient routing protocol for wireless networks,”*ACM Mobile Networks Applications, J. Special Issue on Routing in Mobile Communication Networks*, 1996.

[4] V.D. Park and M. S. Corson, “A Highly adaptive distributed routing algorithm for mobile wireless networks,” in *Proceedings of IEEE INFOCOM'97*, pp. 1405-1413.

[5] D.B. Johnson, D.A. Maltz, and Y. Hu, “The dynamic source routing protocol for mobile ad hoc networks (DSR),” IETF Draft, July 2004.

[6] C-K. Toh, “Associativity-Based Routing for Ad Hoc Mobile Networks,” *International Journal on Wireless Personal Communications*, vol. 4 no. 2, pp. 103-139, March 1997.

[7] Z.J. Haas and M. R. Pearlman, “The performance of query control schemes for the zone routing protocol,” in *Proceedings of. IEEE Symp. Computers and Communication*, 1998.

[8] L. Hanzo and R. Tafazolli, “A Survey of QoS Routing Solutions for Mobile Ad Hoc Networks,” *IEEE Communications Surveys*, vol. 9, no. 2, 2007.

[9] C-Y. Hsu, J.-L. C. Wu and S-T. Wang, “Finding Stable Routes in Mobile Ad hoc Networks,” in *Proceedings of the 18th International Conference on Advanced Information Networking and Application, (AINA'04*), vol. 2, pp. 424-427, March 2004.

[10] W. Su, S. Lee, and M. Gerla, “Mobility Prediction and Routing in Ad Hoc Wireless Networks,” *International Journal of Network Management, Wiley & Sons*, vol.11, no.1, pp: 3–30, 2001.

[11] K. N. Sridhar, & M.C. Chan, “Stability and Hop-count based approach for route computation in MANET,” in *Proceedings of 14th International conference on computer communications and networks (ICCCN'05)*, San Diego, USA.

[12] N-C. Wang, Y-F. Huang and J-C. Chen, “A stable weight-based on- demand routing protocol for mobile ad hoc networks,” *International Journal of Information Sciences*, vol.177, no. 24, pp. 5522–5537, Dec. 2007.

[13] P. Yang and B. Huang, “QoS Routing Protocol Based on Link Stability with Dynamic Delay Prediction in MANET,” IEEE, Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol.1, pp. 515-518, 2008.

[14] M. Hashem and M. Hamdy, “Modified distributed quality-of-service routing in wireless mobile Ad-hoc networks,” in Proceedings of MELECON 2002, pp. 368-378, 2002.

[15] N. Sarma and S. Nandi, “Route Stability Based QoS Routing in Mobile Ad Hoc Networks,” in Wireless Personal Multimedia Communications (WPMC 2007), pp. 203-224, 2007.

[16] Q. Xue and A. Ganz, “ Ad hoc QoS on-demand routing (AQOR) in mobile ad hoc networks”, *Journal of Parallel and Distributed Computing*, vol. 63, no. 2, pp.154–165.

[17] D.S. Thenmozhi, M. Rajaram, “Contention Aware Multi-hop Stable Routing to Provide Quality of Service Based on Multiple Constraints in Mobile Ad Hoc Networks,**”** European Journal of Scientific Research, vol.48, no.4, pp.567-579, 2011.

Wireless Sensor Network Security Challenges and Attacks: A Review

Navjot Sidhu[1], Monika Sachdeva[2] and Krishan Kumar[3]
[1]*Department of Computer Engineering,*
Punjab Technical University, Jalandhar, Punjab
[2,3]*Department of Computer Science & Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab
E-mail: [1]navjotsidhu8@gmail.com, [2]monika.sal@rediffmail.com, [3]k.salujasbs@gmail.com

***Abstract*—Wireless Sensor Networks are a special type of Ad-hoc networks. Although these networks are quite popular now-a-days but limited computing power, energy constraints and security are major challenges for these networks. This paper presents a review on Wireless Sensor Networks and their key challenges. A detailed review of various vulnerabilities and security attacks is also presented. Finally a layer-wise classification of these attacks is also summarized.**

Keywords: Wireless Sensor Networks, Security, Attacks, Protocols, Vulnerabilities

I. Introduction

Recent advances in micro-electro-mechanical systems (MEMS) technology, wireless communications, and digital electronics have enabled the development of low-cost, low-power, and multifunctional sensor nodes [1]. These tiny nodes consist of sensing, data processing, and communicating components. A sensor network is composed of a large number of sensor nodes, which are densely deployed. The sensor network possesses a self organizing capabilities and processing abilities. These features ensure a wide range of applications for sensor networks. Some of the application areas of sensor networks include health e.g. tracking and monitoring doctors and patients, and drug administration in hospitals, military applications, e.g., for battlefield surveillance, reconnaissance of opposing forces and terrain and battle damage assessment, security applications and some other commercial applications like environment control, managing inventory control and vehicle tracking and detection etc.

Although Wireless Sensor Networks are a part of Traditional Ad-hoc Networks. Both of these networks share some common features as Self-organization, energy efficiency, wireless multi-hop. But still there are some key differences between sensor networks and ad hoc networks. Some of the important differences are outlined below [1]–[2]:

Sensor nodes are limited in computation, memory, power resources, and communication speed or bandwidth as compared with ad hoc nodes:

1. The number of sensor nodes in a sensor network can be several times more than the nodes in an ad hoc network.
2. The Wireless Sensor Network has one base station, which has more computing capabilities and act as the controller of the network.
3. Sensor nodes are densely deployed as compared to ad-hoc nodes.
4. Sensor nodes are prone to failures due to various environmental conditions.
5. The topology of a sensor network changes very frequently due to the node failure, joining or mobility.
6. Wireless Sensor Networks are much application specific as compared to ad-hoc networks.

II. Sensor Node and its Construction

A Sensor node in Wireless Sensor Network is a node that is capable of gathering, processing the sensory information and also communicating that information with other nodes connected to it in a network. Huge variations, in the design of sensor devices, are being available. Most sensor devices must have the following hardware:

1. A micro-controller for computation,
2. A small RAM for dynamic data,
3. One or more flash memories to hold the program code and long-lived data,
4. A wireless transceiver,
5. An antenna,
6. An analog-to-digital converter,
7. One or more sensors,
8. And a power source

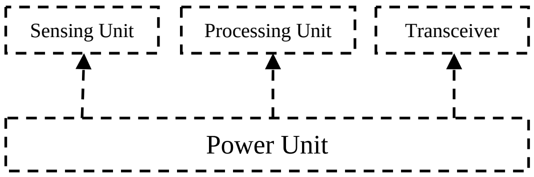


Fig. 1 Components of a Sensor Node

In the end, a sensor node is made up of four basic components: a sensing unit, a processing unit, a transceiver unit and a power unit.

Sensing Unit is composed of two sub-units: Sensors and analog-to-digital converters. Sensors sense the phenomenon and send the observed analog signals to analog-to-digital converters to convert signals in digital form and further send them to processing unit. The processing unit has a small storage unit and manages the procedures to process the sensed information from one or more nodes. A transceiver connects the node to the network. The most important component of sensor node is Power unit which may be supported by some powerful scavenging unit. There can be some other application dependent subunits like location finding system to find the location, mobilize to move sensor nodes when it is required [1].

III. Challenges of Wireless Sensor Network

In sensor networks, wireless nodes self organizes themselves with a dynamic topology. As the number of nodes in a typical sensor network is much higher and to ensure coverage and connectivity, dense deployments are often desired. Sensor nodes have very limited energy, which make them prone to failure. Ideally, sensor network should be power-efficient, small, inexpensive and reliable. According to [3] the key challenges of wireless sensor network are:

1. *Lifetime:* Lifetime is an extremely critical factor and its limiting factor is the energy consumption of sensor nodes. Energy consumption could be reduced by considering the interdependence between individual layers in the network protocol stack.
2. *Flexibility:* Sensor networks should be scalable, i.e. they should be able to adapt dynamic changes to the network. E.g. sensor nodes should adapt changes in the topology, due to failure of some nodes.
3. *Maintenance:* Maintenance in a sensor network leads to complete or partial update of the sensor node program.
4. *Data Collection:* For data collection, sensor network can use ubiquitous mobile agents that randomly move and gather data from sensor nodes and access points. As all data are relayed to a base station, but this form of data collection may affect the lifetime of the network. So, an interesting solution is clustering, which divides network into many clusters. In each cluster, a cluster head collects data from other node and transmits this data to other clusters. The main objective of this technique is to extend the lifetime of the network by limiting number of communications.
5. *Power:* Sensor network use tiny sensors with low computing power, which make them incapable to use complex algorithms. If a sensor node has o do many calculations, its responsiveness will significantly deteriorates.

III. Routing and Data Transmission Mechanism

The sensor nodes usually scattered in field with the capabilities to collect data and send to the sink i.e., a base station. Various protocols are used by sink and other sensor nodes in order to communicate successfully with each other. The protocol stack of wireless sensor networks consists of Application Layer, Transport Layer, Network Layer, Data Link Layer and Physical layer [1]. Each Layer performs a variety of functions, so that data transmission, between tiny sensor nodes, becomes possible. Table 1 represents the different layers of the protocol stack, their functions and protocols used at each layer.

TABLE 1 PROTOCOLS USED AT DIFFERENT LAYERS

LAYER	TASK PERFORMED	PROTOCOLS USED
Application Layer	• Use various application software for Sensing	• Sensor Management Protocol (SMP) • Task Assignment and Data advertising Protocol (TADAP) • Sensor query and data dissemination protocol (SQDDP)
Transport Layer	• Maintain the flow of data	• User Datagram Protocol (UDP) • Transmission Control Protocol (TCP)
Network Layer	• Provide Route to the data supplied by the transport layer • Provide internetworking	• Small Minimum Energy Communication Network Protocol (SMECN) • Sensor protocols for information via negotiation (SPIN) • Sequential assignment routing (SAR) • Low-energy adaptive clustering hierarchy (LEACH)
Data Link Layer	• Provide power awareness and minimize collision	• Media access Control (MAC)
Physical Layer	• Provide robust modulation, transmission and receiving techniques	• Use schemes for frequency selection, carrier frequency generation, signal detection, modulation and data encryption

IV. Vulnerabilities and Attacks

Due to various key challenges like limiting energy, low power, lifetime etc., wireless sensor networks are vulnerable to many threats. Most of these attacks affect limiting energy of sensor networks [4].

A. Classification of WSN Attacks

Generally, attacks are classified as either passive or active depending upon the action they perform. According to [5] attack can be defined as an action to get illegal access to a service, information or to integrity, confidentiality, or availability of a system. In case of wireless sensor networks, attack can be one of the following types:

1. *Passive*: An attack that does not modify data only monitors the communication and threatens the confidentiality.
2. *Active*: An attack that modify and delete existing data and threatens the confidentiality, authentication and data integrity.
3. *Insider*: Steal key information and run malicious code by compromising authorized or legitimate nodes of the network.
4. *Outsider*: Attacker has no particular access to the network.
5. *Mote-Class*: Attacker has access to the minority nodes with similar capabilities.
6. *Laptop-Class*: Attacker has access to powerful devices such as laptop, capable processors, greater battery power and high power antenna.

B. *Principles of WSN Security*

Security principles of wireless sensor networks can be classified as [5]:

1. *Authentication*: defines that data is originated from the authorized source.
2. *Confidentiality*: defines that only authorized sensor nodes can access the messages.
3. *Integrity*: defines that any message has not been modified during transmission by unauthorized node.
4. *Availability*: defines that services provided by wireless sensor network or by a single sensor node must be available whenever necessary.
5. *Data Freshness*: defines that no old data have been used.

C. *Current WSN Attacks*

Wireless sensor networks are vulnerable to security attacks due to the broadcast nature of transmission, limiting energy or nodes are often placed in a dangerous environment. In many applications, the data obtained by the sensing nodes needs to be kept confidential. In the absence of security measures a false or malicious node could intercept private information or send false information to sensor nodes in the network [6]. The brief overview of current wireless sensor network attacks is given below [4], [6], [7]:

1. Eavesdropping: It is a passive attack which only listen the network to intercept information, but does not modify data. That’s why, it is very difficult to detect.
2. *Radio Jamming*: An attacker sends the radio waves at the same frequency that is used by other authorized sensor nodes of the network.
3. *Message Injection*: It is an active attack, in which aim of the attacker is to send the false messages on the network to corrupt the records or to saturate the network.
4. *Message Replication*: It is also an active attack, here attacker catches the transmitted packets over the network and sends those packets to wrong nodes of the network.
5. *Node Destruction*: It is a type of physical attack in which ne or many nodes of sensor network are destroyed, making network not to work to destroy a node the link two nodes. In this type of attack, the attacker can also reprogram the sensor nodes.
6. *Denial of service*: This is another active attack which makes the wireless sensor network out of order by sending large amounts of data to the sensors to be active and consumes their energy.
7. *Hello Flooding*: With an attack of Hello flooding, an attacker can use a device with large enough transmission power for compromising every node in its neighbour.
8. *Black Hole Attack*: In black hole attack, at first a malicious node is inserted into the network. This malicious node changes the routing tables of the network. The aim is to force a maximum of neighbouring nodes to send data to it. Once it captures all sent data, it does not forward or replies back.
9. *Gray Hole Attack*: It is a variant of the black hole attack. In this attack the malicious node replays all information concerning the route and non critical data. That’s why this attack is more difficult to detect.
10. *Wormhole Attack*: Unlike the black hole attack, this attack needs to insert in the network at least two malicious nodes. These nodes are connected by powerful connection. This attack wrongs the other nodes of the network and proposes a quicker path. Nodes choose this shortest path to send their data, and in actual they send their data to malicious nodes.
11. *Sinkhole Attack*: In this attack the malicious node attacks directly the data, which circulate near the sink i.e. base station. To perform this attack, the malicious node offers the quickest path to reach the sink. All nodes, which are near the malicious node, send data for the sink which may be captured by the attacker.
12. *Sybil Attack*: In Sybil attack, a malicious sensor which is masquerading as multiple sensors, modifies the routing table.
13. *Message Alteration*: A malicious node catches a message and changes it, by adding wrong information or deleting some information.
14. *Slowdown*: An attacker can make use of malicious nodes to slow down the network. This attack prevents a sensor to sleep in different ways, in order to consume its battery quickly.

V. Layer-wise Classification of Attacks

A Wireless Sensor network is comprised of a large number of sensors that collaboratively monitor various environments. To collect data from WSNs, base stations and aggregation points are commonly used. As they usually have more resources than normal sensor nodes. Security is one of the most important aspects that deserve great attention [7].

TABLE 2 Layer-wise Classification of Attacks

LAYER	THREAT	COUNTERMEASURES
Application Layer	• Selective Message Forwarding • Data Aggregation Distortion	• Integrity Protection • Confidentiality Protection
Transport Layer	• Flooding	• Manage connection Request
Network Layer	• False Routing • Message Replication • Black Hole • Sink Hole • Selective Forwarding • Worm Hole • DOS • Sybil Attack	• Routing Access Restrictions • False Routing Information Detection • Wormhole Detection
Data Link Layer	• Traffic Manipulation • Identity Spoofing	• Misbehavior Detection • Identity Protection
Physical Layer	• Eavesdropping • Radio Jamming • Node Destruction	• Access Restriction • Encryption

Table 2 presents, a classification of attacks based on the layering model of Open System Interconnection, along with some potential countermeasures [5]–[8].

Because of the nature of wireless communications, resource limitation on sensor nodes, size and density of networks, unknown topology prior to deployment, and high risk of physical attacks, it is a challenge to provide security in Wireless Sensor Networks. The Wireless Sensor Network has general security requirements of availability, integrity, authentication, confidentiality and non-repudiation. These security requirements can be provided by distribution mechanisms with the requirement of scalability, efficiency key connectivity and resilience [8].

VI. Conclusion

Recent micro-electro-mechanical system advances have allowed use multifunctional sensor networks. But information in these networks is still not secure and vulnerable to many attacks. Communications over wireless channels are insecure and easily susceptible to various kinds of attacks. It is impractical to protect each individual sensor node from physical or logical attack. The security and vulnerabilities of a wireless sensor network depend upon the particular application for which sensor network is deployed. Most of the attacks in wireless sensor networks are caused by inserting false or wrong information by malicious nodes within the network.

Presented classifications are crucial for future implementation of Wireless Sensor Networks.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “Wireless sensor networks: a survey”, Elsevier Science Computer Networks, vol. 38, pp. 393–422, 2002.

[2] X. Chen, K. Makki, K. Yen, and N. Pissinou, “Sensor Network Security: A Survey”, IEEE Communications Surveys & Tutorials, Vol. 11, No. 2, pp. 52-73, 2009.

[3] D. Puccinelli and M. Haenggi, “Wireless Sensor Networks: Applications and Challenges of Ubiquitous Sensing”, IEEE Circuits and System Magazine, pp. 19-29, 2005.

[4] D. Martins and H. Guyennet, “Wireless Sensor Network attacks and Security Mechanisms: a Short Survey”, Proceedings of 13th IEEE International Conference on Network-Based Information Systems, pp. 313-320, 2010.

[5] H. Modares, R. Salleh and a. Moravejosharish, “Overview of Security Issues in Wireless Sensor Networks”, Proceedings of 3rd IEEE International Conference on Computing Intelligence, Modelling & Simulation, pp. 308-311, 2011.

[6] K. Sharma and M.K Ghose, “Wireless Sensor Networks: An Overview on its Security Threats”, IJCA Spacial Issue on MANETs, pp. 42-45, 2010.

[7] K. Xing, S.S. R. Srinivasan, M. Rivera, J. Li and X. Cheng, “Attacks and Countermeasures in Sensor Networks: A Survey”, Network Security, Springer, pp. 1-28, 2005.

[8] Z.S. Bojkovic, B.M. Bakmaz and M.R. Bakmaz, “ Security Issues in Wireless Sensor Networks”, International Journal of Communications, Issue 1, vol. 2, pp. 106-115, 2008.

[9] S. Raman, A. Prakash, K.B. Pulla and P. Srivastava, “Wireless sensor networks: A Survey of Intrusion and their Explored Remedies”, International Journal of engineering Science and Technology, vol. 2(5), pp. 962-969, 2010.

[10] A.K. Pathan, H. Lee and C.S. Hong, “Security in Wireless Sensor Networks: Issues and Challenges”, ICACT, pp. 1043-1048, 2006.

[11] J. zheng and A. Jamalipour, “Wireless Sensor Networks-A Netwoking Perspective”, John Wiley, 2009.

[12] H. Suo, J. Wan, L. Huang and C. Zou, “ Issues and Challenges of Wireless Sensor Networks Localization in Emerging Applications”, Proceeding of IEEE International Conference on Computer Science & Electronics Engineering, pp. 447-451, 2012.

[13] G. Padmavthi and D. Shanmugapriya, “A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks,” International Journal of Computer Science and Information Security, vol. 4, pp. 1-9, 2009.

[14] T. Kavitha and D. Sridharan, “Security Vulnerabilities in Wireless Sensor Networks: A Survey”, Journal of Information Assurance and Security 5, pp. 31-44.

[15] X. Chen, K. Makki, K. Yen and N. Pissinou, “Sensor Network Security: A Survey”, IEEE Communication Surveys & Tutorials, vol. 11, no. 2, pp. 52-73.

[16] C. Karlof and D. Wagner, “Secure Routing in Wireless Sensor Networks: Attacks and Countermeasures”, Elsevier Ad-hoc networks, pp. 293-315, 2003.

A New Robust and Secure Approach to SVD-3 Level DWT Video Watermarking

Pawandeep Kaur[1] and Sonika Jindal[2]
[1,2]Department of Computer Sc. & Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, India
E-mail: [1]pawandeep234@gmail.com, [2]sonikamanoj@gmail.com

***Abstract*—Digital watermarking is used to protect digital content such as images, audio and videos that have been tampered maliciously and with higher accuracy. Digital video watermarking is a new and merging area of research to exploit different ways in order to prohibit illegal replication and exploitation of digital contents. In this paper, to maintain the quality of video and to ensure the ownership we propose a new SVD-3 Level DWT watermarking embedding technique. Singular value decomposition (SVD) is an important transform technique in robust digital watermarking. We apply the 3 level DWT and SVD on selected frames and embed the watermark into randomly selected frames with the help of secret key to authenticate the video by considering the video quality, robustness and video imperceptibility.**

Keywords: Digital Video Watermarking, Secret Key, Scaling Factor, 3 Level DWT SVD Algorithm

I. INTRODUCTION

In the past several years a rapid growth in multimedia (audios, videos, images) and illegal transfer of this multimedia content over the internet are becoming important issues in digital era. This leads the development of new technologies providing security to this multimedia content. Digital watermarking is used to protect this sensitive information using different watermarking technologies. Video watermarking is relatively a new technique in multimedia technology. [1] Video watermarking is the process in which watermark is embedded in a video sequence by using a secret key. The amount of information that can be embedded in the video sequence is called payload. The extraction is performed at the other end using the same secret key as shown in Fig. 1**.** The embedded watermark should be robust against variety of attacks such as Subtractive attacks, Distortive attacks, Additive attacks, Filtering, Cropping, Compression, Rotation and Scaling attacks, so that video can be protected from illegal copying and provide security against several other attacks that only performed on videos such as frame dropping, frame swapping and frame averaging [2].The two types of watermark can be used such as visible watermark and invisible watermark. We can add the watermarks either in the whole frames of video or in certain frames depending upon the requirement [3].

[4] Video watermarking is very different from image watermarking, even though some techniques can be viewed as an extension to it. [1] [4] video watermarking is mainly used in two domains: spatial domain, frequency domain. The first category is spatial domain watermarking in which watermark is embedded in frames by directly modifying the pixel values of that frame or replacing the bits of selected frame pixels [4]. In second category [4] Frequency domain watermarking techniques, first coefficients of transformed video frames are modified and then transformations are applied and at last the inverse process is applied to get the watermarked video. Discrete Fourier transforms (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) [5, 6, and 7] and the Singular Value Decomposition (SVD) [8, 9] is common transforms for watermarking. Watermarking is mostly used in frequency domain because of human visual system is more sensitive to low frequency coefficients and less sensitive to high frequency coefficients. [10] Depending upon the various applications, video watermarking is used in fingerprinting, copyright protection, video authentication, copy control and broadcast monitoring. Apart from these applications video watermarking systems has some properties including effectiveness, data payload, blind or informed detection, false positive rate, capacity, robustness, perceptual transparency, security, cost, sensitivity, and scalability [2].

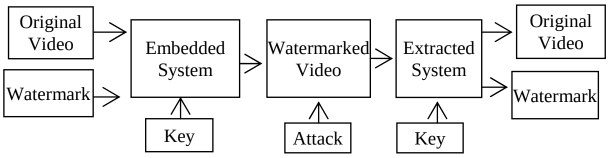


Fig. 1 A General Video Watermarking Process

The rest of the paper is organized as section 2 describes Related work. Section 3 describes Proposed Architecture. Section 4 describes Proposed A lgorithm. Section 5 defines Experimental Results. Section 6 demonstrates conclusion.

II. RELATED WORK

DWT: [11] [14] It divides an image into two sections such as in lower resolutions as well as in higher resolutions. Lower resolution means LL components and higher resolution means horizontal (HL), vertical (LH) and diagonal (HH) detail components. The low frequency part is further divided into two sections of high and low frequencies. This process is repeated number of times to compute multiple scale wavelet decomposition. [12] Proposed a method in which 3D DWT is applied using perceptual mask and embedding is performed by weighing the mark through the defined mask and then the Inverse 3D DWT (IDWT) is performed:

- *Advantages*: More accurate model because its properties similar to HVS and more robust to noise addition.
- *Disadvantages*: Higher frequencies change the quality of image.

SVD: It is a mathematical tool which decomposes a matrix into two orthogonal matrices and one diagonal matrix consisting of the singular values of the matrix [13].The SVD mathematical technique provides an elegant way for extracting algebraic features from an image and improves watermark robustness and resistance against many kinds of attacks. SVD is a useful method to separate the system into a set of linearly independent components. A digital Image X of size MxN can be represented by its SVD as follows:

$$X = USV^T \quad (1)$$

$U = [U_1, U_{22}.....U_m]$,

$V = [V_1, V_{22}.... V_n]$,

$$S = \begin{bmatrix} \sigma_1 & & \\ & 0 & \\ & & \sigma_2 \end{bmatrix} \quad (2)$$

SVD is more applicable in watermarking because of following reasons:

- SVD is able to efficiently represent the intrinsic algebraic properties of an image, where singular values correspond to the brightness of the image.
- Singular values have good stability, which means a small perturbation added to an image will not significantly change the corresponding singular values. [18]

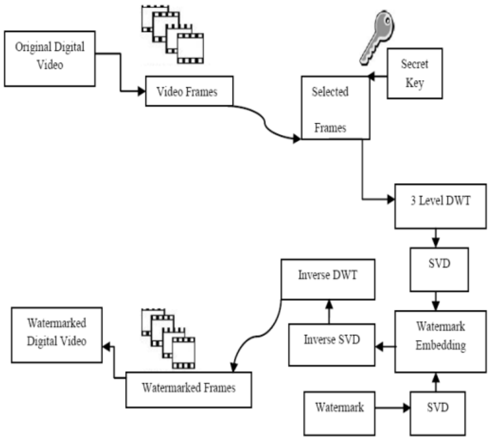


Fig. 2 Procedure of Watermark Embedding

III. Proposed Architecture

The proposed method effectively hides the secret data into video using existing video watermarking techniques. Fig. 2 give a complete overview of data flow in proposed algorithm. This method uses some frames of video to hide the secret data. The frames selected to hide secret data are random frames and not sequential frames. Hence each frame that contains the secret data can be identified using secret key, a 10 digit number provided by user. The selection of frames is done by using several functions that are made up from secret key. So, watermark is embedded in whole video and not in some parts of video. We also set up a passkey identifier to give only four trials to the user and if the user inserts more than 4 wrong keys then it means he/she is trying to find out the watermarked frames by trying random keys. If four wrong entries are made by user then the video will be damaged leaving no data behind.

IV. Proposed algorithm

In this section, we have discussed some motivating factors in the design of our approach to video watermarking. We have used DWT and SVD for developing the algorithm. Among various tools, SVD and DWT are more reliable in digital watermarking. Due to the fact of localization in both spatial and frequency domain, wavelet transform is the most preferable transform among all other transforms. After converting the video into frames, we have applied 3 levels DWT on selected frames. In the next stage, the SVD is applied to selected sub-bands and embed the same original watermark by modifying the singular values. Embedded watermark in middle frequencies increases the robustness to variety of attacks. The procedure of embedding a digital watermark into the original video is depicted in Fig. 2. At last, inverse SVD and inverse DWT is applied in order to reconstruct the watermarked digital video. After getting the watermarked video the extraction process is performed at other end in order to check the extracted watermark resembles with original one or not.

A. *Watermark Embedding Algorithm*

- Apply DWT to the selected frames repeatedly up to the third level.
- Perform SVD transform on approximation and all the detail parts in third level of wavelet transform, $f_Q = U_Q S_Q V^T$ Where Q ∈ {LL3, LH3, HL3, HH3}.

$$f_Q = U_Q\, S_Q\, V^T \quad (1)$$

- Perform SVD transform on watermark,

$$W = U_W S_W V^T{}_W \quad (2)$$

- In general, embedded watermark at this stage. Modify the singular values of approximation and all the detail parts with the singular values of the watermark as:

$$\gamma_Q^* = \gamma_Q + \alpha_Q \gamma_W \quad (3)$$

- Here, is scale factor of combined transform, which value is 0.04.
- Take inverse combined transform and reconstruct the watermarked video.

B. *Watermark Extraction Algorithm*

- Apply DWT to selected watermarked frames repeatedly up to the third level.
- Apply SVD transformation on approximation and all details parts up to the third level of wavelet transform, Where Q ∈ {LL3, LH3, HL3, and HH3} and get the combined transform coefficient γ_Q^*
- Extract singular values of watermark from approximation and all detail parts.

$$\gamma_{W^*}^{Q} = \frac{\gamma_Q^* - \gamma_Q}{\alpha_Q} \qquad (4)$$

- Extract the watermark from video frames.

$$W_Q^* = U_W S_Q^* V_W^T \qquad (5)$$

- After detecting all estimates of watermark, sum up all these estimates and normalized $\overline{W_Q^*}$ between [0, 1].
- Reproduced the watermark,

$$W_Q^* = \sum_{i=1}^{Q} w_Q^* \qquad (6)$$

V. Experimental Results

The main focus of this algorithm is its dynamic and key dependent frame selection technique [3]. We have implemented and experiment it using MATLAB. The experimental results are as below which show original frames and corresponding watermarked frames. We test the proposed watermarking algorithm with different variations using colored host video clips. Each video clip is partitioned into different number of frames. We employed "McD" video sequence in AVI format where total number of frames we calculated is 926 and frame rate 25 fps. We have selected 10 random frames and embed watermark such as "jeep.jpg" of size (512 × 512) in that frames as shown in Fig. 4. The 10 random original frames are shown in Fig. 3 and their corresponding watermarked frames are shown in Fig. 5. Watermarked Video quality was estimated by SSIM, PSNR, BER and MSE.

Fig. 3 Original Frames

Fig. 4 Watermark Image

Fig. 5 Watermarked Frames

TABLE 1 Calculated Value of SSIM, PSNR, MSE and BER for Scaling Factor 0.04

Video	Frame no.	SSIM	PSNR	BER	MSE
McD	20	0.99804	52.679	0.018983	0.350
	119	0.99777	53.162	0.018988	0.313
	211	0.99739	53.1703	0.018807	0.313
	299	0.99658	52.7252	0.018966	0.347
	406	0.99638	52.3973	0.019085	0.374
	502	0.99650	52.6176	0.019005	0.355
	570	0.99651	52.91	0.0189	0.332
	658	0.99659	53.5664	0.018668	0.286
	767	0.99712	53.3088	0.018759	0.303
	838	0.99871	52.7933	0.01894	0.341

We then tested the robustness and quality of watermarked video using a scaling factor 0.04 and different performance evaluation metrics. For each frame we have calculated the SSIM, PSNR, MSE and BER as shown in above Table I.

A. *To Check the Robustness*

The PSNR is a quality metric used to determine the degradation in the embedded image with respect to the host image or also defined as ratio between maximum power of a signal and power of distorted signal [16]. It is most easily defined via the mean squared error (MSE) as:

$$\text{PSNR} = 10log_{10}\frac{L*L}{MSE}$$

The MSE [16] defined it as average squared difference between a reference image and a distorted image. It is calculated as:

$$MSE = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j=1}^{Y}(c(i,j) - e(i,j))^2$$

The BER [16] defined it as the ratio that describes how many bits received in error over the number of the total bits received. It is often expressed as percentage and calculated by comparing bit values of embedded image and cover image.

BER = $P/(H * W)$

The SSIM is a method for measuring the similarity between two images. SSIM is designed to improve on traditional methods like peak signal to noise ratio (PSNR) and mean squared error (MSE), which have proven to be inconsistent with human eye perception. It is calculated by formula given below:

$$\text{SSIM}(x,y)=\frac{(2\mu_x\mu_y+c1)(2\sigma_{xy}+c2)}{(\mu_x^2+\mu_y^2+c1)(\sigma_x^2+\sigma_y^2+c2)}$$

The value calculated shows that propose DWT-SVD based video watermarking algorithm is imperceptible. The calculated PSNR value is 57.2156 db which shows quality of watermarked video appear visually identical to the original one and there is no degradation in visual quality. The value calculated for SSIM is 0.99871 which shows the structural similarity between original video and watermarked video.

In order to check the quality of extracted watermark, the normalized Cross-correlation (NC) value between the original watermark and extracted watermark is calculated for different frames using scaling factor 0.04, which is defined as:

$$\text{NC}=\frac{\sum_{i=0}^{M_1}\sum_{j=0}^{M_2}[W(i,j)W'(i,j)]}{\sum_{i=0}^{M_1}\sum_{j=0}^{M_2}[W(i,j)]^2}$$

Where W and W` represent the original image and extracted watermark image, respectively. The watermark extraction using scaling factor 0.04 is shown in Fig. 6 which show that correlation value of extracted watermark is near to 1 and extracted watermark is same as original one.

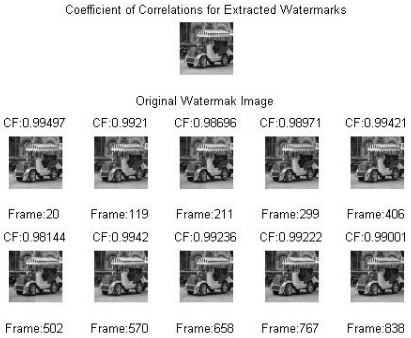


Fig. 6 Watermark Extraction with Scaling Factor 0.04

B. *Choice of Approppriate Scaling Factor*

It is actually a hard step for choosing the suitable scaling factor. Usually, the scaling factor is chosen to be a scalar value. In most of literature the scaling factor is chosen between 0 and 1[17].Table II shows the SSIM, PSNR, MSE,BER of the watermarked video and the correlation coefficient (NC) of the extracted watermark for several scaling factors. From this table, the higher scaling factor is, the worse the robustness and invisibility of watermark will be.

Figure 7, Fig. 8 and Fig. 9 show the original watermark and the extracted watermark for different scaling factors, i.e $\alpha = 0.9, 0.5$ and 0.1.

TABLE 2 AVERAGE VALUES OF SSIM, PSNR, BER AND MSE FOR WATERMARKED VIDEO AND EXTRACTED WATERMARK USING VARIOUS SCALING FACTORS

Different Parameters of Watermarked Video and Normalized Cross-Correlation Values for Extracted Watermark						
Video	**Scale Factor**	**SSIM**	**PSNR**	**BER**	**MSE**	**NC**
McD	0.9	0.9151	38.557	0.0259	9.0844	0.92415
	0.5	0.9539	40.509	0.0246	5.801	0.95862
	0.1	0.9960	50.122	0.1995	0.6343	0.98845
	0.04	0.9992	57.215	0.0174	0.1235	0.99082

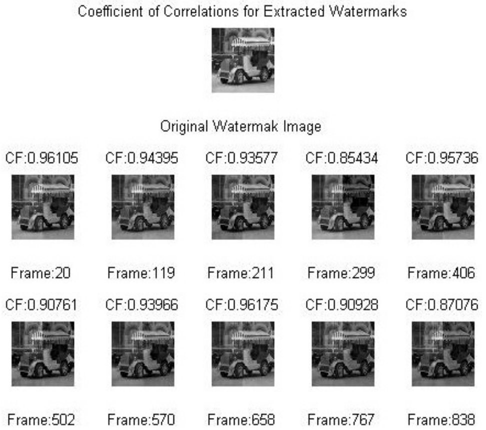


Fig. 7 Extracted Watermark Using $\alpha = 0.9$

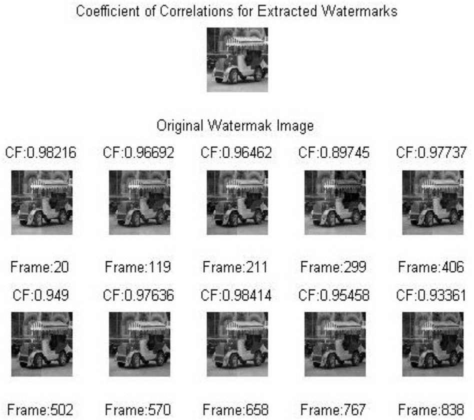


Fig. 8 Extracted Watermark Using $\alpha = 0.5$

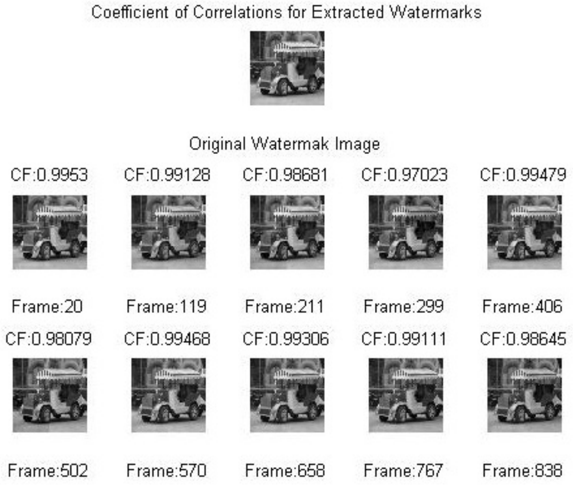


Fig. 9 Extracted Watermark Using $\alpha = 0.1$

VI. Conclusion

This paper provides that proposed video watermarking SVD and 3 Level DWT algorithms is dynamic and key dependent which provides the security against video authentication. The selection of frames is done using a secret key which provide the secure selection of random frames and the extracted watermark resembles with the original one. The choice of appropriate scaling factor gives the robustness and good quality watermarked video.

Acknowledgment

This is to express my sincere gratitude to Dr. Satvir Singh, Associate Professor, Department of Electronic & Communication Engineering, SBS State Technical Campus, Ferozepur (Punjab), India, for sparking in me the enthusiasm and initiative to discover and learn. I am truly thankful to him for guiding me through the entire paper and being as a motivator in this learning curve.

References

[1] Jayamalar, T and Radha, V, "Survey on digital video watermarking techniques and attacks on watermarks," International Journal of Engineering Science and Technology, vol. 2, Pp. 6963-6967, 2010.

[2] Potdar, Vidyasagar M and Han, Song and Chang, Elizabeth, "A survey of digital image watermarking techniques," Industrial Informatics, 2005.

[3] Madia, Jigar and Dave, Kapil and Sampat, Vivek and Toprani, Parag, "Video Watermarking using Dynamic Frame Selection".

[4] Doerr, Gwena and Dugelay, Jean-Luc, "A guide tour of video watermarking," Signal processing: Image communication, Elsevier, vol. 18, Pp.263-282,2003.

[5] Tay P, Havlicek JP. "Image watermarking using wavelets", Pp. 258–261, 2002.

[6] Kundur D, Hatzinakos D., "Digital watermarking using multi-resolution wavelet decomposition".Int Conf Acoust Speech Signal Proc, Pp. 2969–72, 1998.

[7] Wu C, Zhu W-P Swamy MNS., "A watermark embedding scheme in wavelet transform domain". In: IEEE Region 10 Conference Proceedings: Analog and Digital Techniques in Electrical Engineering, vol. A, Pp.279–82, 2004.

[8] Loukhaoukha K, " Chouinard J-Y, "Hybrid watermarking algorithm based on SVD and lifting wavelet transform for ownership verification." In: IEEE. Pp.177–82, 2009.

[9] Gorodetski V, Popyack L, Samoilov V, Skormin V. "SVD based approach to transparent embedding data into digital images," In: Proc. International Workshop on Mathematical Methods, Models and Architectures for Computer Network Security (MMM-ACNS'01). 2001.

[10] Lu, Gaoyan and Zhang, Yongping and Liang, Fengmei and Zheng, Dechun,"Survey of Video Watermarking" Video Engineering,vol.21, Pp.009,2012

[11] Sinha, Sanjana and Bardhan, Prajnat and Pramanick, Swarnali and Jagatramka, Ankul and Kole, Dipak K and Chakraborty, Aruna," Digital video watermarking using discrete wavelet transform and principal component analysis," International Journal of Wisdom Based Computing,vol.1, Pp 7--12, 2011.

[12] Campisi, Patrizio and Neri, Alessandro," Video watermarking in the 3D-DWT domain using perceptual masking," IEEE, vol.1, Pp.I--997, 2005.

[13] K.-L. Chung, W.-N. Yang, Y.-H. Huang, S.-T. Wu, Y.-C. Hsu, "On svd-based watermarking algorithm," Applied Mathematics and Computation Pp 54–57, 2007.

[14] Preda, Radu O and Vizireanu, Dragos N," A robust digital watermarking scheme for video copyright protection in the wavelet domain," Measurement, Elsevier,vol. 43, Pp 1720—1726,2010.

[15] Rastegar, Saeed and Namazi, Fateme and Yaghmaie, Khashayar and Aliabadian, Amir," Hybrid watermarking algorithm based on Singular Value Decomposition and Radon transform,"AEU-International Journal of Electronics and Communications, vol. 65, Pp 658—663,2011.

[16] A. K. Singh, N. Sharma, M. Dave, A. Mohan, "A novel technique for digital image watermarking in spatial domain," in: Parallel Distributed and Grid Computing (PDGC), 2012 2nd IEEE International Conference on, IEEE, pp. 497–501, 2012.

[17] Mohammad, Ahmad A and Alhaj, Ali and Shaltaf, Sameer, "An improved SVD-based watermarking scheme for protecting rightful ownership," Signal Processing Elsevier, vol. 88, Pp 2158—2180,2008.

[18] Rajab, Lama and Al-Khatib, Tahani and Al-Haj, Ali, "Hybrid DWT-SVD video watermarking," Innovations in Information Technology, 2008. IIT 2008. International Conference on, IEEE,, Pp 588--592,2008.

A Review: IPTV over WiMAX Networks

Rajdeep Kaur[1] and Seema Baghla[2]
[1,2]Department of Computer Engineering,
Yadavindra College of Engineering, Talwandi Sabo, Punjab, India
E-mail: [1]er.rajdeepvirk@gmail.com, [2]garg_seema238@yahoo.co.in

***Abstract*—Worldwide Interoperability for Microwave Access (WiMAX) technology is the wireless system capable of offering high QoS at high data rates for IP networks. Deployment of Video on Demand (VoD) over the next generation (WiMAX) has become one of the passionate subjects in the research these days, and is expected to be the most revenue generator in the coming years. The objective of this paper to provide a survey of IPTV services over WiMAX Networks, the key success factors of IPTV over WiMAX and to provide comparison of WiMAX networks over other networks like WiFi.**

Keywords: IPTV, QoS, VoD, WiFi, WiMAX

I. Introduction

WiMAX is basically a new shorthand term for IEEE Standard 802.16, which was designed to support the European standards. 802.16 predecessors (like 802.11a) were not very accommodative of the European standards. The IEEE wireless standard has a range of up to 30 miles, and can deliver broadband at around 75 megabits per second. This is theoretically, 20 times faster than a commercially available wireless broadband [1]. WiMAX can be used for wireless networking like the popular WiFi. WiMAX, a second-generation protocol, allows higher data rates over longer distances, efficient use of bandwidth, and avoids interference almost to a minimum. WiMAX can be termed partially a successor to the Wi-Fi protocol, which is measured in feet, and works, over shorter distances [1].

WiMAX have generated interest among researchers these years because of their potential usage in wide variety of applications [4]. WiMAX supports a variety of modulation and coding schemes and allows the scheme to change on a burst-by-burst basis per link, depending on channel conditions [6]. The bandwidth and range of WiMAX make it suitable for the following potential applications:

1. Providing portable mobile broadband connectivity across cities and countries through a variety of devices.
2. Providing a wireless alternative to cable and digital subscriber line (DSL) for "last mile" broadband access.
3. Providing data, telecommunications (VoIP) and IPTV services (triple play).
4. Providing a source of Internet connectivity as part of a business continuity plan.
5. Smart grids and metering.

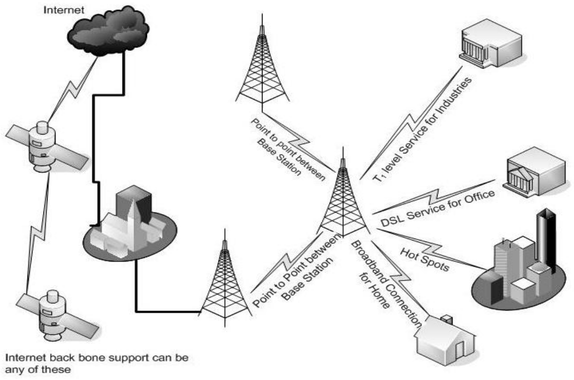

Internet
Point to point between Base Station
T_1 level Service for Industries
DSL Service for Office
Hot Spots
Point to Point between Base Station
Broadband Connection for Home
Internet back bone support can be any of these

Fig. 1 WIMAX Network [8]

This paper is divided into five sections. The section II consists of IEEE 802.16 protocol architecture, section III is of IPTV, section IV explains IPTV services and WiMAX and section V concludes the paper with a conclusion.

II. IEEE 802.16 Protocol Architecture

The IEEE 802.16 protocol architecture is structured into two main layers: the Medium Access Control (MAC) layer and the Physical (PHY) layer. MAC layer consists of three sub-layers. The first sub-layer is the Service Specific Convergence Sub-layer (CS), which maps higher level data services to MAC layer service flow and connections. The second sub-layer is Common Part Sub-layer (CPS), which is the core of the standard and is tightly integrated with the security sub-layer. This layer defines the rules and mechanisms for system access, bandwidth allocation and connection management. The MAC protocol data units are constructed in this sub-layer. The last sub-layer of MAC layer is the Security Sub-layer which lies between the MAC CPS and the PHY layer, addressing the authentication, key establishment and exchange, encryption and decryption of data exchanged between MAC and PHY layers [2].

The PHY layer provides a two-way mapping between MAC protocol data units and the PHY layer frames received and transmitted through coding and modulation of radio frequency signals.

Comparisons and confusion between WiMAX and Wi-Fi are frequent because both are related to wireless connectivity and Internet access.

1. WiMAX is a long range system, covering many kilometres that uses licensed or unlicensed spectrum to deliver connection to a network, in most cases the Internet.

2. Wi-Fi uses unlicensed spectrum to provide access to a local network.
3. Wi-Fi is more popular in end user devices.
4. Wi-Fi runs on the Media Access Control's CSMA/CA protocol, which is connectionless and contention based, whereas WiMAX runs a connection-oriented MAC.
5. WiMAX and Wi-Fi have quite different quality of service (QoS) mechanisms:

5.1. WiMAX uses a QoS mechanism based on connections between the base station and the user device. Each connection is based on specific scheduling algorithms.

5.2. Wi-Fi uses contention access-all subscriber stations that wish to pass data through a wireless access point (AP) are competing for the AP's attention on a random interrupt basis. This can cause subscriber stations distant from the AP to be repeatedly interrupted by closer stations, greatly reducing their throughput.

6. Both 802.11 (which include Wi-Fi) and 802.16 (which include WiMAX) define Peer-to-Peer (P2P) and ad hoc networks, where an end user communicates to users or servers on another Local Area Network (LAN) using its access point or base station. However, 802.11 supports also direct ad ho or peer to peer networking between end user devices without an access point while 802.16 end user devices must be in range of the base station. [11]

III. Internet Protocol Television (IPTV)

Internet Protocol Television (IPTV) has become popular as it promises to deliver the content to users whenever they want. IPTV is a set of multimedia services that are distributed throughout an IP network, where end of user receives video streams through a set-top-box (STB) connected to a broadband connection. IPTV is often combined with the services of VoD. VoD services contents are not live but pre-encoded contents available at any time from servers. These services must possess an adequate level of quality of service, security, interactivity, and reliability. [12] From the perspective of the provider, IPTV includes the video acquisition, video processed and video secure distribution on the IP network infrastructure. One official definition approved by the International Telecommunication Union focus group on IPTV (ITU-T FG IPTV) is:

"IPTV is defined as multimedia services such as television/video/audio/text/graphics/data delivered over IP based networks managed to provide the required level of quality of service and experience, security, interactivity and reliability".

Another more detailed definition of IPTV is the one given by Alliance for Telecommunications Industry Solutions (ATIS) IPTV Exploratory Group in 2005:

"IPTV is defined as the secure and reliable delivery to subscribers of entertainment video and related services. These services may include for example, Live TV, Video on Demand (VOD) and Interactive TV (iTV). These services are delivered across an access agnostic, packet switched network that employs the IP protocol to transport the audio, video and control signals [9]. In contrast to video over the public Internet, with IPTV deployments, network security and performance are tightly managed to ensure a superior entertainment experience, resulting in a compelling business environment for content providers, advertisers and customers alike."

IPTV services may be classified into three main groups:

1. Live television, with or without interactivity related to the current TV show.
2. Time-shifted television: catch-up TV (replays a TV show that was broadcast hours or days ago), start-over TV (replays the current TV show from its beginning).
3. Video on demand (VOD): browse a catalogue of videos, not related to TV programming.

IPTV is distinguished from Internet television by its on-going standardization process (e.g., European Telecommunications Standards Institute) and preferential deployment scenarios in subscriber-based telecommunications networks with high-speed access channels into end-user premises via set-top boxes or other customer-premises equipment.

IV. IPTV Services & wimax

WiMAX technology is one of the access technologies that enable transmission of IPTV Services. Transmitting IPTV over WiMAX aims to make IPTV services available to users anywhere, anytime and on any device. The QoS for delivering IPTV services depends especially on network performance and bandwidth [7]. Scalable Video Coding (SVC) has achieved significant improvements in coding efficiency with an increased degree of supported scalability relative to the scalable profiles of prior video coding standards.[5] Transmitting SVC encoded videos over WiMAX networks is an effective solution which solves many of the video transmission problems over these networks.

In general, IPTV services can be classified by their type of content and services [13] [14]:

On-demand content: With pre-encoded and compressed content, a customer is allowed to browse an online movie catalogue, to watch trailers, and to select a movie of interest. Unlike the case of live video, a customer can request or stop the video content at anytime and is not bound by a particular TV schedule. The playout of the selected movie starts nearly instantaneously on the customer's TV or PC.

Live content: In this case, a customer is required to access a particular channel for the content at a specific time, similar to accessing a conventional TV channel. A customer cannot request to watch the content from the beginning if he or she joins the channel late. Similar to a live satellite broadcast, live content over IPTV can be a showing of a live event or a show encoded in real-time from a remote location, such as a soccer game.

Managed services: Video content can be offered by the phone companies who operate the IPTV business or obtained from syndicated content providers, in which the content is usually well-managed in terms of the coding and playout quality, as well as in the selection of video titles. Bandwidth for delivery and customer equipment are arranged carefully for serving the best play out performance and quality to the customers.

Unmanaged services: The technology of IPTV itself enables play out of any live or on demand video content from any third party over the Internet. Therefore, nothing stops a customer from accessing video content directly from any third party online such as YouTube (or Google Video), individuals, or an organization.

V. The Key Success Factors

A. *Economy of Scale*

Economy of scale characterizes a production process or service operation, in which an increase in the number of producing units may cause a decrease in the average fixed cost of each unit. By optimizing the economy of scale for operating IPTV services, one can minimize the risks and secure the early advent of ultimate success. This translates to the need of an access network technology that can support more subscribers and mobile TV for future requirements [14].

B. *Scheduled Live Content and Quality Assurance*

Quality of service and quality of experience for end users have been identified as critical requirements of IPTV services[10]. In the long run, watching IPTV content will be just like surfing different Web sites over the Internet. Watching unmanaged live or on-demand content offered by different service and media providers in the world would provide the true value of IPTV services to customers. However, an IPTV channel is still critical to ensure comparable TV quality and experiences similar to those of the conventional cable, satellite, or digital TV services. Offering managed and scheduled SDTV programs with a quality guarantee is required to secure a head start and the success of IPTV service [14].

VI. Conclusion and Future Work

This paper has provided us with information related to IPTV over WiMAX along with WiMAX characteristics, IPTV services and key success factors of IPTV with WiMAX. In this paper, we have discussed the major problems that are faced in delivering the good quality of service over WiMAX. This paper can act as a base for those who are new to this field. IPTV & WiMAX can be used jointly to provide better quality of on demand audio & video services over the network.

In WiMAX, nodes are free to move without disconnectivity with the network. So due to the movement of nodes there is possibility that misbehaviour nodes come across the network. So it is necessary to analyse the effect of misbehaviour nodes in WiMAX. In the near future, we will try to find the impact of IPTV over WiMAX with misbehaviour nodes so that quality of service of IPTV over WiMAX can be improved.

References

[1] S.S. Riaz Ahamed, D.Mahesh "Performance analysis and special issues of broadband strategies in the computer communication" Journal of Engineering Science and Technology, Vol.1 (2), 2009, 73-89.

[2] Eugen Borcoci, "WiMAX Technologies: Architectures, Protocols, Resource Management and Applications" IARIA CTRQ Conference June 29-July 5, 2008.

[3] Md. Ashraful Islam, Riaz Uddin Mondal and Md. Zahid Hasan "Performance evaluation of WiMAX physical layer under adaptive modulation techniques and communication channels", (IJCSIS), Vol. 5 No.1,pp. 111-114, 2009.

[4] M.A. Mohamed, M.A. Abou-Elsoud and H.F. Mohamed "Effect of WiMAX networks on IPTV technology", International Journal of Computer Science Issues(IJCSI), Vol. 10 Issue 6 No 2, pp. 235-243,November 2013.

[5] Jamil Hamodi, Khaled Salah and Ravindra Thool "Evaluating the performance of IPTV over fixed WiMAX", International Journal of Computer Applications (0975 – 8887) Vol. 84 No 6,pp. 35-43,December 2013.

[6] Jamil Hamodi, Ravindra Thool, Khaled Salah, Anwar Alsagaf and Yousef Holba "Performance study of mobile TV over mobile WiMAX considering different modulation and coding techniques," Int. J. Communications, Network and System Sciences, 2014, 7, 10-21 Published Online, January 2014.

[7] K. Ain, M. Tarafder, Sh. Khan and Md. Ali, "Path loss compensation technique for WiMAX technology based communication system," International Journal of Engineering Science and Technology, 2011.

[8] I. Uilecan, C. Zhou, and G. Atkin, "Framework for delivering IPTV services over WiMAX wireless networks," Proc. IEEE EIT 2007, Chicago, IL, Vol.4, No. 6, pp. 470–475, May 2007.

[9] Will Hrudey and Ljiljana Trajkovic, "Streaming video content over IEEE 802.16/WiMAX broadband access," OPNETWORK 2008, Washington, DC, Vol.6, no.2, pp. 469-470, Aug. 2008.

[10] A. Shehu, A. Maraj, and R.M. Mitrushi, "Analysis of QoS requirements for delivering IPTV over WiMAX technology," International Conference on Software, Telecommunications and Computer Networks (SoftCOM), vol.2, no.1, pp. 380-385, 2010.

[11] WiMAX forum, WiMAX Technology info, www.quantum WiMAX.com.

[12] Jamil M. Hamodi and Ravindra C. Thool, "Investigate the performance evaluation of IPTV over WiMAX networks" International Journal of Computer Networks & Communications (IJCNC) Vol.5, No.1, January 2013.

[13] M.Dies – L.Zuccaro, "IPTV over WiMAX" Network Infrastructures A.A. 2008-2009.

[14] James she,Fen Hou and Pin HAn-Hou, "IPTV over WiMAX: Key Success Factors, Challenges an solutions" IEEE Communications Magazine 09/2007.

Classification of Software Projects Based on Software Metrics: A Review

Kunal Chopra[1], Monika Sachdeva[2] and Sunil Dhawan[3]
[1]Department of Computer Science & Engineering,
Shaheed Bhagat Singh State Technical Campus, Firozpur, India
[2,3]Department of Computer Science,
NIMS University, Jaipur, Rajasthan, India
E-mail: [1]kunalraichopra@gmail.com, [2]monika.sal@rediffmail.com,
[3]sunildhawan007@gmail.com

***Abstract*—Software metrics are developed and used by the many software organizations for the evaluation and confirmation of good software code, working and maintenance of the software product. Software metrics measure and identify various types of software complexities such as size metrics, control flow metrics and data flow metrics. Such observed and calculated software complexities should be continuously calculated, understood and controlled. One of the significant objective of software metrics is that it is applicable to both a process and product metrics. It is always a clear fact that the high degree of complexity in modules is bad in comparison to a low degree of complexity in modules. Metrics can be used in different phases of the software development lifecycle. This paper reviews the theory, called "software complexity metrics for evaluation of software projects", and analysis can be done based on various static and dynamic parameters. We will try to evaluate and analyze various aspects of software metrics in determining the quality and improvise the working of the software product development.**

Keywords: Software Metrics, Lines of Code, Control Flow Metrics, Structural Testing

I. Introduction

"Metrics don't solve problem, people solve problem, Metrics provide information so that people can solve problems."[1]

METRICES: Metrics was introduced in the context of software measurement which has become an essential tool for the good software engineering. So Metric can be defined as a quantitative measure of degree to which a system, component or process possesses a given attribute. "A handle or guess about a given attribute" for example "Number of errors found per person hours expended."

A. *Need and Importance of Software Metrics*

Why do we measure???...

- Determine the quality of the current product or process.
- Predict qualities of a product/process.
- Improve quality of a product/process.

Many of the best software developers measure the characteristics of the software to get some sense that whether the requirements are consistent or complete, whether the design is of high quality whether the code is ready to be tested. So basically software engineers measures the attributes of the product so as to estimate the project will be ready for the delivery or the budget may perhaps be exceeded. [2]

Measurement is a necessary practice for understanding, improving and controlling our environment. It requires rigor and care and though it has huge impact on software engineering. Measurement need is directly related to goals we set and question we ask while developing a software.

Thus Measure can defined as ***a*** quantitative indication of extent, amount, dimension, capacity, or size of some attribute of a product or process for example Number of errors in the entire code as it said that, "When you can measure the product you are talking numbers that means we can easily count or express an entity in figures similarly but when we cannot measure the product or an entity we cannot express it into numbers our knowledge is unsatisfactory type it may perhaps be the beginning of the knowledge."

B. *Need of Measurement in Everyday Life*

Measurement is essential in our daily lives and measuring has become the necessity and well accepted fact. It exists in the heart of many systems that governs our lives.

Economic measurements determines price and pay increases. In Medical sciences measurements help doctors to diagnose certain illness. Measurements related to atmospheric systems are the basis for whether forecasting. Without measurement, technology cannot operate. Measurement perhaps may not solely the domain of professional technologists. Everyone of us uses it in our daily life.

For instance, while making a trip we do measure distance, selects our route, measure our speed and predict the time when we arrive at our destination. So measurement helps us to understand our world, interact with our surroundings and improve our lives.

C. *What is Measurement?*

Measurement can be well predicted with a example in a shop if we compare the price of one commodity with another. In a garment shop we contrast sizes. And in case of journey we can compare distance travelled to distance remained. So we can make the calculations and predictions accurate according to well defined set of rules. [3]

So measurement may be defined as the process by which the symbols and numbers are assigned to attributes of entities in the real world in such a way so as to describe them according to clearly defined set of rules.

The measurement provides information about attributes of entities. An entity can be an object or an event in real world. So an entity can be well descrived by identifying characteristics that are important to us in distinguishing one entity from another. An attribute is the feature or property of entity.

So when we describe these entities by using their attributes, we define those attributes by using numbers and symbols. Thus the price is designated as the rupees or dollars sterling, while height is defined in terms of inches and feet. Those numbers and symbols are abstractions that we usually use to reflect people's perception in the real world.

Thus measurement can be determined as the process whose definition is not accurate. To understand what measurement is we may have to ask host of questions which may difficult to answer.

1. Height of a person is commonly known attribute that can be measured. But other attributes of people, such as intelligence creates a fuss.
2. Height is commonly measured in terms of meters, inches and feet. These different scales measure the same attribute. But we can also measure height in terms of miles and more appropriate for the measurement of distance of satellite above earth but not for the measurement of the height of the person which again makes measurement definition far from accurate.
3. The accuracy of the measurement depends upon the measuring instrument and the definition of the measurement. For example length can measured with accuracy as long as the ruler is accurate and used in proper way.
4. Once we attain measurements for different aspects of real world, we need to analyze them and define conclusions about the entities from which they were derived. It also requires that what sort of changes or manipulations can we apply for the results of measurement? For example why it is acceptable to say that Joe is twice as tall as Fred but not acceptable to say that it is twice as hot today as it was yesterday?

D. Making things Measurable

"What is not measurable, make it measurable."

The above stanza suggests that one of the aims of science is to find ways to measure attributes of the things in which we are interested. Measurement makes concepts more clear and therefore more understandable and controllable. Thus, as scientists, we should find out ways to measure world; where we can already measure, we can make our measurement better.

To improve the implementation of measurement in software engineering, we need not to restrict type of measurements we make. Really measuring the un-measurable should improve our understanding of particular entities and attributes, and making software engineering as powerful as other engineering disciplines.

Strictly speaking, there are two kind of quantification: measurement and calculation.

Measurement is termed as direct quantification, as in the measurement of the height of a tree or the weight of a shipment of bricks. On the other hand calculation is indirect quantification, where we take measurements and combine them into a quantified item that describe the attribute whose value is to be determined.

For instance, when a city inspectors assign a valuation to a house, they calculate it by using certain formula that combines variety of factors which includes number of rooms, the overall floor space and the type of heating and cooling. Thu the valuation is termed to be quantification, not a measurement, and it expression as a number makes it more useful than qualitative assessment alone.

So eventually, it is necessary to modify our surrounding or our practices in order to measure something new or in an innovative way. It can be achieved by using a new tool, adding a new steps in a process, or using a new method. In many cases, change is difficult for people to accept, there are management issues to be considered whenever a management program is implemented or changed.

E. Measurement in Software Engineering

We have seen the importance of measurement in our daily life, measurement has become an essential and well accepted attribute of life. In this section, we will see instances of software engineering to see why measurement is needed.

Software Engineering briefs the collection of techniques that apply an engineering approach to the construction and maintenance of software products. It includes activities like managing, costing, planning, modeling, analyzing, specifying, implementing, testing and maintaining.

In engineering we try to impend each activity to be well understood and maintained so that there are fewer surprises as the software is designed, specified, built and maintained. On the another hand computer science gives the theoretical foundations for building software, software engineering focuses on implementing the software in a controlled and specific manner.

The significance of software engineering cannot be understood, since software pervades our lives. From banking transactions to air traffic control, from oven controls to air bags,, and sophisticated power plants to sophisticated weapons, our life and quality of life depends upon software.

In software engineering we use various software models and theories for example in making an electrical circuit we appeal to theories like Ohms Law which gives the relation between resistance, current and voltage in the circuit. Once the scientific method suggests the validity of the subject concern, the measurement or the truth of the story, we continue to use measurement to apply the theory to practice. Thus to build a circuit with a specific current and resistance, we know what voltage is required and we use instruments to measure that we have such voltage in the given battery.

It is difficult to predict the mechanical, electrical and civil engineering without a central for measurement. Indeed science and engineering can neither be effective nor practical without measurement. But measurement in software engineering has been considered a luxury. For most development projects:

1. Gilbs principle of Fuzzy Targets: projects without clear goals will not achieve their goals clearly. For example we promise to make a reliable, user-friendly and maintainable without specifying clearly and objectively what these terms mean.
2. We do not quantify or predict the quality of products we produce. Thus we cannot tell a potential user how reliable the product will be in terms of likelihood of failure in a given period of use, or how much work will be needed to port the product to a different machine environment.

Since measurements are made they are often done inconsistently, infrequently, inconsistently and incompletely. The incompleteness can be frustrating to those who really want to use the results. For instance, a developer may claim that 80% of all software costs involve maintenance, or that there on average 55 faults in every 1000 lines of software code. But we are not always told how these results were obtained, how experiments were designed and executed, which entities were measured and how, and what were realistic error margins. Without this additional information, we remain skeptical and unable to decide whether to apply results to our own situations. [4]

The Software complexity is based on well-known software metrics, this would be likely to reduce the time spent and cost estimation in the testing phase of the software development life cycle (SDLC), which can only be used after program coding is done. Improving quality of software is quantitative measure of the quality of source code.

This can be achieved through definition of metrics, values can be calculated by analyzing source code or program is coded. A number of software metrics widely used in the software industry are still not well understood.

Although some software complexity measures were proposed over thirty years ago and some others proposed later. Sometimes software growth is usually considered in terms of complexity of source code.

Various metrics are used, which unable to compare approaches and results. In addition, it is not possible or equally easy to evaluate for a given source code.

Software complexity, deals with how difficult a program is to comprehend and work with Software maintainability, is the degree to which characteristics that hamper software maintenance are present and determined by software complexity.

II. The Scope of Software Metrics

Software metrics include many activities that may include some sort of measurement. It may help in determining various activities like:

1. Structural and complexity metrics.
2. Management by metrics.
3. Evaluation of methods and tools.
4. Cost and effort estimation.
5. Productivity measures and models.
6. Data collection.
7. Reliability models.
8. Quality modals and measures.
9. Performance evaluation and models.

A. Why do We Need to Classify

From software engineering point of view software development experience shows, that it is difficult to set measurable targets when developing software products.

Produced/developed software has to be testable, reliable and maintainable. On the other side, “You cannot control what you cannot measure”. [5]

In software engineering field during software process, developers do not know if what they are developing is correct and guidance are needed to help them accustom more improvement. Software metrics are facilitating to track software enhancement. Various industries dedicated to develop software, and use software metrics in a regular basis. Some of them have produced their own standards of software measurement, so the use of software metrics is totally depending upon industry to industry. In this regards, what to measure is classified into two categories, such that software process or software product.

But ultimately, main goal of this measure is customer satisfaction not only at delivery, but through the whole development process. [6]

Various software metrics have been discovered and proposed by the researchers if we take a glimpse of the history of software metrics. The software metrics range through size, design and complexities proposed by

McCabb (1976), Helstead (1977), Lorenz (1993) and Chidamber and Kermer (1994) were chosen for the improvisation in design and development of the software projects. The discovered metrics domains were non OO and OO designing in software engineering which were implemented empirically onto various software projects so as to increase the productivity and quality of the project.

Huge budget is being spent in the maintenance and improving the quality of software projects based on the criteria set by the proposed metrics. But this mechanism is somewhat not good in essence that these approaches are implemented during the maintenance phase or rarely at the design phase. This can be prevented if we classify our software projects in accordance with the software metrics.

Based on non-OO and OO design metrics we can broadly classify our software projects in the following category:

a. Size based projects.
b. Design oriented projects.
c. Approach based projects.
d. Program weakness.
e. Failures.
f. Functionality.
g. Complexity.
h. Dependency.

III. Types of Sotware Metrics

As we have discussed earlier that first obligation of any software measurement activity is identifying the attributes and entities we wish to measure. In software there two such classes:

- **Processes Metrics** are collection of software related activities.
- **Products Metrics** documents or deliverables that result from a process activity.

A. Software Process Metrics

Software process metrics involves measuring of properties of the development process and also known as management metrics. These metrics include the cost, effort, reuse, methodology, and advancement metrics. Also determine the size, time and number of errors found during testing phase of the SDLC.

B. Software Product Metrics

Software process metrics involves measuring the properties of the software and also known as quality metrics. These metrics include the reliability, usability, functionality, performance, efficacy, portability, reusability, cost, size, complexity, and style metrics. These metrics measure the complexity of the software design, size or documentation created.

C. Size Metrics: Lines of Code

Certain size metrics were proposed for measuring the software like LOC(Lines of Code), KLOC (1000 Lines of Code), SLOC(Statement Lines of Code). Lines of code is actually count of instruction statements. It's count is usually for executable statements. [7] Since the LOC count gives the program size and complexity, it is not a surprise that the more lines of code there are in a program, the more defects are expected. More surprisingly, researchers found that defect density(defects per KLOC) is also significantly related to LOC count. Previous studies pointed to a negative relationship: the larger the module size, the smaller the defect rate. For example, Basili and Perricone (1984) examined FORTRAN modules with fewer than 200 lines of code for the most part and found higher defect density in the smaller modules. Shen and colleagues (1985) studied software written in Pascal, PL/S and Assembly language and found an inverse relationship existed upto about 500 lines. Since larger modules are generally more complex, a lower defect rate is somewhat counterintuitive.

D. Halstead Complexity Metric (1977)

It distinguishes software science from computer science. According to computer science a computer program is a collection of tokens that can be classified as either operators of operands. [9] The primitive measures of Halstead's software science are:

n1 = Number of distinct operators in a program
n2 = Number of distinct operands in a program
N1 = Number of operator occurrences
N2 = Number of operand occurrences

Given the attribute measures based on that, Halstead developed a system of equations which expresses the overall program length, the potential minimum volume for an algorithm, total vocabulary, the, the actual volume(the number of bits required to specify a program), the program level (a measure of software complexity), program difficulty, and other features such as development effort and projected number of faults in the software. Halstead major equations include the following:

a. Program Length (N) = $N1+N2$
b. Program Vocabulary (n) = $n1+n2$
c. Volume of a Program (V) = N*log2n
d. Potential Volume of a Program (V*)=(2+n2)log2(2+n2)
e. Program Level (L) = L=V*/V
f. Program Difficulty (D) = 1/L
g. Estimated Program Length (N) = n1log2n1+n2log2n2
h. Estimated Program Level (L) = 2n2/(n1N2)
i. Estimated Difficulty (D) = 1/L = n1N2/2n2

j. Effort (E) = V/L = $V*D$ = ($n1$ x $N2$) / $2n2$
k. Time (T) = E/S

[“*S*” is Stroud number (given by John Stroud), the constant “*S*” represents the speed of a programmer. The value “*S*” is 18]

One major weakness of this complexity is that they do notmeasure control flow complexity and difficult to compute during fast and easy computation.

E. McCabe Cyclomatic Complexity by (1976)

It was designed to indicate a programs testability and understandability. It is the classical graph theory cyclomatic number, indicating the number of regions in the graph. As applied to the software, it is the number of linearly independent paths that comprise the program. The M is equal to the number of binary decisions plus 1. [8]

If all the decisions are not binary, a three way decision can be counted as two binary decisions and n-way case statement is counted as n-1 binary decisions. The cyclomatic complexity metric is additive. The complexities of several graphs considered as a group is equal to the sum of individual graphs complexities. The general formula to compute the cyclomatic complexity is:

M= V(G) = e – n + 2p where.
V(G) = Cyclomatic number of G.
e = Number of edges.
n = Number of nodes.
p = Number of unconnected parts of the graph.

We can compute the number of binary node (predicate), by the following equation.

$V(G)= p+1$
where, $V(G)$= Cyclomatic Complexity
P= number of nodes or predicates.

The problem with McCabb’s Complexity is that, it fails to distinguish between different conditional statements (control flow structures). Also does not consider nesting level ofvarious control flow structures.

F. Design Metrics

In 1994 Chidamber and Kermer proposed six OO design and complexity metrics, which became the commonly referred to CK metric suite:

1. *Weighted Method per Class* (WMC): WMC is the sum of the complexities of the methods, whereas complexity is measured by cyclomatic complexity. If one consider all the methods of a class to be of equal complexity, then WMC is simply the number of methods defined in each class. And the average of WMC is the average number of methods per class. [10]
2. *Depth of Inheritance tree* (DIT): This is the length of the maximum path of class hierarchy from the node to the root of the inheritance tree. [10]
3. *Number of Children of Class* (NOC): This is the number of immediate successors (subclasses) of class in a hierarchy.[10]
4. *Coupling between object classes* (CBO): An object class is coupled with another one if it invokes another one’s member functions or instance variables. CBO is the number of classes to which a given class is coupled. [10]
5. *Response for Class* (RFC): This is the number of methods that can be executed in response to a message received by an object of that class. The larger the number of methods that can be invoked from a class through messages, the greater the complexity of the class. It captures the size of the response set of a class. The response set of a class is all the methods called by the local methods. RFC is the number of local methods plus the number of methods called by the local methods. [10]
6. *Lack of Cohesion on Methods* (LCOM): The cohesion of a class is indicated by how closely the local methods are related to the local instance variables in the class. High cohesion indicates good class subdivision. The LCOM metric measures the dissimilarity of methods in a class by the usage instance variables. LCOM is measured as the number of disjoint sets of local methods. Lack of cohesion increases complexity and opportunities for error during the development process. [10]

IV. Conclusion

Software metric are used in analyzing and maintaining the quality of the software development process and it is one of the most important process associated with SDLC. We can use the metrics to analyze various factors that impact the design and then the performance of the software product. Thus the metrics we have reviewed become an integral part of the process known as software development. The deployment of the metrics are indeed a very big task and it provides a vast array of opportunities for the programmers to refer this as a document for referencing for classifying metrics.

REFERENCES

[1] Singh Yogesh & Pradeep Bhatia, “ Module Weakness—A New Measure”, ACM SIGSOFT Software Engineering Notes,81,July,1998.

[2] Norman E. Fenton & Shari Lawrence Pfleeger "Software Metrics A Rigorous and Practical Approach " PWS Publishing Company, 2-1, 1997.

[3] Martin Neh, " Software Metrics for Product Assessment", McGraw Hill Book Co., UK, 2003.

[4] Henry S. & Kafura D., "Software Structure Metrics Based on Information Flow", IEEE Trans. On Software Engineering SE-7, 5, 510-518, Sept. 1981.

[5] Paul Goodman, " Practical Implementation of Software Metrics", McGraw Hill Book Co., UK, 1993.

[6] Halstead M.H., "Elements of Software Science", New York, Elsevier North Holland, 1977.

[7] Mrinal Kanti Debbarma, Swapan Debbarma, Nikhil Debbarma, Kunal Chakma, and Anupam Jamatia, "A Review and Analysis of Software Complexity Metrics inStructural Testing, March 2013.

[8] Dhawan Sunil, Wadhwa Manoj, Identification of Software Metrics for Software Projects, IJACEN, 23-27,2013.

[9] Stephen H. Kan, "Metrics and Models In Software Quality Engineering", Second Edition Pearson, 2-82 2002.

[10] Shyam R. Chidamber and Chris F. Kermer, "A Metrics Suite For Object Oriented Design", June 1994.

A Review of Handoff and Location Management Techniques in Ad-hoc Networks

Aanchal Bawa[1] and Jyoteesh Malhotra[2]
[1,2]*Department of Computer Science,*
GNDU, Regional Campus, Jalandhar, Punjab
E-mail: [1]aanchalbawacse@gmail.com, [2]jyoteesh@gmail.com

***Abstract*—With the invention of Web and wireless mobile communications and the increasing number of mobile subscribers, mobility management came to be as the most important and challenging problems for wireless mobile communication over the Web. Mobility management enables the existing networks to detect a mobile subscriber's point of attachment for sending the data to required terminal (i.e., location management), and maintain a mobile subscriber's connection as it continues to change its point of attachment (i.e. handoff management). The paper concentrates on the issues and functionalities of location management and handoff management in terms of mobile adhoc networks.**

Keywords: Mobility Management, Location Management, Multi-hop Nature, Handoff Management

INTRODUCTION

With the growing era of mobiles in adhoc networks, mobility management is one of the important area need to be covered. A wireless ad-hoc network is a decentralized type of wireless network. The network is ad hoc because it does not depends on a pre existing network, for example routers in wired networks or access points in managed (infrastructure) wireless networks. Rather, each node participates in routing by passing data packets to another node, so as to judge the dynamic connectivity between the nodes in the network. [6] Mobility management is one of the major functions of a GSM or a UMTS network that allows mobile phones to work. The aim of mobility management is to track where the subscribers are, allowing calls, SMS and other mobile phone services to be delivered to them.

A Mobile Ad-hoc Wireless Network (MANET) is a self-configuring network of mobile nodes. Nodes serve as routers and may move arbitrarily. There is no static infrastructure and the communication network must be able to adapt to changes because of movement and other dynamics. Most of the MANET protocols do not assume that position data is available. However, if such position data is available then efficient location based communication protocols are applicable. The main problem in MANET is to find a multi-hop route between the source and the target of information. It is clear that if all the intermediate router nodes are moving that this type of network is very much affected by mobility. Especially if one takes into account that the transmitting range is rather restricted to a limited supply of energy. The main mobility problems for a MANET are routing a message, multicasting a message, and upholding the network routing tables for these issues. [1]

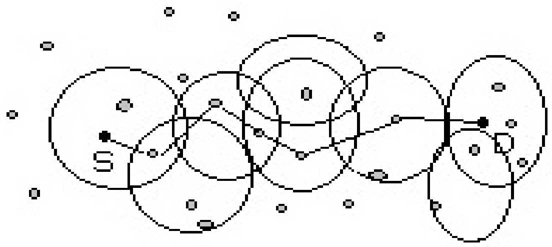


Fig. 1 Multi hop Packet Delivery of a MANET [6]

A. *Mobility Management Techniques*

Mobility management consists of handoff management and location management. In handoff management, the mobility is handled in such a way that during IP subnet handovers the current application connections remain intact, thus ongoing connections will be either preserved or restarted after a movement.

B. *Location Management*

Location management means the terminal also has to inform the current location (current IP on an interface) to its communication peers or some intermediate router in the network. This way all willing communication nodes can access a moving between subnets. Therefore, location management consists of two tasks:

1. Location Update – To track the location of mobile nodes, their location must be registered and this registration must be updated on every change.
2. Traffic delivery – Using the location information, traffic is delivered(routed) to the mobile node's current location.[2]

C. *Handoff Management*

With the moving nature of mobiles need to switch the base stations frequently to remain in the calling process and this process is known as handoff management. Thus, handoff management is the process through which a mobile node keeps itself in connection while it moves from one access point to another. Fig. 2 shows how a node can be disconnected if handoff is not implemented and thus calling process is crashed. To eliminate this problem handoff management is required. Handoff process consists of three stages. Firstly the initialization is being done either by mobile, dynamic network modifications or network agent. In second stage, new connection is being established by searching for the resources for handoff management and required additional routing operations are being performed. Finally, taking care of Quality of Service (QOS) the data flow control maintains delivery of data from previous connection path to recent one.

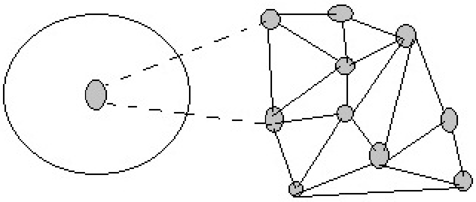


Fig. 2: Disconnection of Node Near the MANET Edge. The Node has Moved out of Range and Can no Longer Reach the Rest of the Network. [6]

Depending on the movement of the mobile device, handoff can be classified in various types. In a broad sense, handoffs may be of two types:

1. intra-system handoff (horizontal handoff).
2. inter-system handoff (vertical handoff). [3]

A homogenous network usually supports intra-system handoff. This type of handoff occurs when the signal strength of the serving Base Station goes below a certain threshold value.

Handoffs can also be classified into two type namely hard handoff and soft handoff.

A **hard handoff** can also be said as "break before make" connection. The BS handoffs the MS's call to another call under the control of MSC and then drop the call. The link of previous BS is terminated before transferring the connection of mobile node to new BS; the MS is linked to no more than one BS at any given time and this type of handoff is called hard handoff. Hard handoff is supported mainly by TDMA (Time division multiple access) and FDMA (Frequency division multiple access), where adjacent channels uses different frequencies in order to reduce channel interference. So when the MS moves from previous BS to next BS, it becomes difficult for it to communicate with both BSs. [4]

A **soft handoff** can maintain multiple connections with neighboring cells. It is used where cells use same frequency and can be accessed by different code words such as CDMA (Code division multiple access). Each MS maintains an active set where BSs are added when the RSS exceeds a given threshold and removed when RSS drops below another threshold value for a given amount of time specified by a timer. When a presence or absence of a BS to the active set is encountered soft handoff occurs. [5]

II. RELATED WORK

As the paper specifies, there exists the various issues in mobility management of adhoc networks such as location management, handoff etc. Thus to eliminate each issue it is required to know all the properties of the network specified. There are various location management strategies even such as Power-up and power-down location update, Time-based, Distance-based, Zone-based, Parameter-based, ordered update and implicit location update. Choosing the best combination of routing protocol and location update strategy will lead to a network with no disconnection of nodes. With this it must be checked that the network is FDMA, TDMA or CDMA as each access technique support different handoff management strategy, the first two works on Hard Handoff and former one is on Soft Handoff. Thus, to have a reliable network, it must be analyzed properly so as to know the best mobility management techniques.

Lidong Zhou, Zygmunt J. Haas (1999) [7] states that adhoc networks are not only used for military purposes but also for commercial use and thus security constraints are very important. The paper concentrates on various security issues and different solutions that can be applied such as cryptography schemes such as threshold cryptography.

Qing-An Zeng and Dharma P. Agrawal (2001) [4] says mobility is the most important feature of a wireless cellular communication system. To continue service of mobile stations handoff is very necessary otherwise it may lead to breakage of connection. The paper contains the different types of handoffs and how they affect handoff.

Brent Ishibashi, Raouf Boutaba (2004) [6] states that a highly dynamic topology is required in the mobile connection establishment and maintenance as links between nodes are created and broken, as the nodes are dynamic in the network. The network multihop nature not only affects the source or destination's mobility but intermediate nodes also. Thus a topology is the extreme requirement so as to have an efficient adhoc network and thus the resulting routes can be extremely volatile. To have better understanding of adhoc networks number of factors are need to be studied such as link, routes and the environment, several parameters are required to be studied such as radio dimension range, number of nodes, network dimensions and mobility parameters are also examined so as to gain maximum speed and less wait time. Thus several properties require to be considered so as to establish a reliable adhoc network and a optimized MANET protocol.

Nasıf Ekiz, Tara Salih, Sibel Küçüköner and Kemal Fidanboylu in 2005 [5] discussed that the quality of a cellular communication is measured by continuation of the call and this can be achieved by handoff process which enables a call to be transferred from one base station to another and thus maintains it and quality too. In this paper, an overview is represented regarding the issues in handoff intiation and maintenance and discussed about the different handoff techniques present.

Christian Schindelhauer (2007) [1] surveyed mobility patterns and mobility models for wireless networks. Mobility patterns are aerial, robot, dynamic medium and outer space motion. This paper presents the characteristics of each and shortly mentions the specific problems. It presents the specifics of cellular networks, mobile ad hoc networks, and sensor networks regarding mobility. It also discusses about the research regarding mobility in wireless networks and specifies mobility models from literature.

D. LI, J. WANG, L. ZHANG, H. LI and J. ZHOU (2009) [9] A MANET is a self configuring network so it is difficult to design a location management scheme that is both scalable and cost-efficient. A corporative location management scheme is introduced in this paper called CooLMS for MANETs. CooLMS combines the strength of grid based location management and pointer forwarding strategy to achieve high scalability and low signaling cost.

Jaydip Sen(2010) [3] states that as the demand of reliable network is increasing day by day so we need to an appropriate network supporting different types of traffic and obstacles coming in the network and the different Quality of Service. Different users require different types of services and there also exists different wireless technologies which satisfy these needs but as these wireless networks act as complementary to each other as some networks supports one application and some another and if they are integrated together they will result into a best network according to the user's requirements. To have the best network a proper handoff scheme is required. There also exist different handoff schemes in a heterogeneous networking environment which are also presented in the paper.

Umang, B.V.R. Reddy, M.N. Hoda (2011) [8] stated that as the nodes remain in moving state due to dynamic nature of adhoc network, they must be monitored on regular basis. The paper specifies the importance of management schemes in adhoc networks. Mobility and Traffic pattern of mobility models are generated by using AnSim Simulator and related with real life.

Anju Gill (2012) [10] provides an overview of routing protocol, traffic types, underlying issues and challenges related to security, mobility and resource limitation and also give possible solution for them.

Jani Puttonen (2013) [2] had discussed in dissertation about mobility management in IP based wireless environments. Mobility management can include both handovers within one technology and selection of access technology in a heterogeneous overlapping environment. Real time information about the link status and quality as well as user preferences is taken into account in the interface selection. The objective is to offer an Always Best Connected access to the user, and seamless handovers.

As per the previous research and discussion it has been concluded that there exists the various issues which are being studied in various papers and solution to them is being researched. Some common issues related to this paper are given in the below table:

Paper	Issue	Solution	Remarks
Lidong Zhouand and Zygmunt Haas	Security Issues like poor physical protection	New cryptographic schemes, such as threshold cryptography can be applied.	Only cryptography cannot provide appropriate security required in the network.
Qing-An Zeng and Dharma P. Agrwal	Forced termination probability of ongoing calls	Traffic models and handoff schemes such as non priority and priority schemes	Termination problems can be solved by choosing accurate traffic model and handoff scheme
Christian Schindelhauer	Find a multi-hop route between the source and the target of information.	Choosing the best routing protocol and nodes positioning.	Sometimes the best protocol and nodes positioning also cannot work and network remains weak.
D.LI,J.WANG, L. ZHANG, H. LI, J. ZHOU	Location management scheme which is scalable and cost-efficient	CooLMS for MANETs is designed to achieve it.	Simulation results show CooLMS performs better than other schemes under certain circumstances.
Jaydip Sen	Delay in movement detection	Handoff for the forward and reverse direction	Handoff resolves this issue.
Umang, B.V. R. Reddy, M.N. Hoda	Nodes remain in dynamic state always	Regular monitoring through mobility management generated by An Sim Simulator	Simulator will help in detecting the position of the node and hence results in a reliable network.

III. Open Issues

Based on the literature survey done in previous sections, the important issues still need attention, said by research committee. To have efficient wireless networks with seamless services, all-IP framework and heterogeneous technologies will be used in future. But to have a reliable network given issues should be examined and needed to be resolved:

a. *Location and handoff management in wireless overlay networks*–Mobility management is the one of the important issue for future as the networks will be inherently hierarchal where different areas are need to be covered.

b. *Quality of Service Issue*-In future networks it is required to provide guaranteed QOS to mobile terminals. Various problems may occur while providing QOS such as inter-system handoff, location management etc.

c. *Security Issues*-Security is the essential part especially where security sensitive applications exist. The various attributes required to be examined are integrity, authentication, availability, confidentiality and non-repudiation.
d. *Limited radio range*-MANETS has limited radio range due to low transmission power.
e. *Mobile node functioning as a router* – Each node has its own running protocol and thus it acts as a router and host too.

IV. Future Scope and Conclusion

About adhoc networks, it a network which does not rely on previous data and hardware such as routers and mobility management is the technique which allows mobile phones to work. It is one of the major issues of adhoc networks which can be further sub divided in Location Management and Handoff Management. Location Management is a method through which location of the mobility node is being detected and updated in the network. It includes location update strategies and traffic delivery constraints. Handoff Management refers to the technique which helps to switch the mobility nodes from one base station to another through two different methods namely hard handoff and soft handoff. Issues in location registration and handoff management have been identified and several existing mechanisms have been presented. Since global roaming will be an increasing trend in future, attention has been paid on mechanisms which are applicable in heterogeneous networks. Media Independent Handover Services of IEEE 802.21 standard as an enabler for handover has also been presented. Mobility management is one of the major issues required to be work in future to have a reliable and secure network.

Acknowledgment

I am thankful to the almighty for the strength provided through which I completed this paper and please allow me to dedicate my acknowledgment of gratitude towards parents, teachers, family, friends, and in essence, all sentient beings.

References

[1] Christian Schindelhauer, “Mobility in Wireless Networks”, 2007
[2] Jani Puttonen,”Mobility Management in Wireless Networks”, 2006.
[3] Jaydip Sen,”Mobility and Handoff Management in Wireless Networks”, 2010.
[4] Qing-An Zeng and Dharma P. Agrwal, Department of Electrical Engineering and Computer Science,University of Cincinnati, “Handoff in Wireless Mobile Networks”,2001.
[5] Nasıf Ekiz, Tara Salih, Sibel Küçüköner and Kemal Fidanboylu, “An Overview of Handoff Techniques in Cellular Networks”,2005.
[6] Brent Ishibashi, Raouf Boutaba, “Topology and mobility considerations in mobile ad hoc networks”, 2004.
[7] Lidong Zhou, Zygmunt J. Haas, “Securing Ad Hoc Networks”, November-1999.
[8] Umang, B.V.R. Reddy and M.N. Hoda. Article: Study of Mobility Management Schemes in Mobile Adhoc Networks. *International Journal of Computer Applications* 17(7):42-47, March 2011.
[9] D. LI, J. WANG, L. ZHANG, H. LI and J. ZHOU, "A Cooperative Location Management Scheme for Mobile Ad Hoc Networks," *Int'l J. of Communications, Network and System Sciences*, Vol. 2 No. 8, 2009, pp. 732-741.
[10] Anju Gill, Chander Diwaker, “Behavioral Study of Issues and Challenges in Mobile Adhoc Network”, Volume 2, Issue 5, May 2012.

A Review of Trends and Opportunities for Data Mining Applications in Telecommunications Industry

Parmpreet Kaur[1] and Jyoteesh Malhotra[2]
[1,2]Department of Computer Sc. & Engineering,
GNDU Regional Campus, Jalandhar, Punjab, India
E-mail: [1]parmpreet89@gmail.com, [2]jyoteesh@gmail.com

***Abstract*—The development of mobile networks and internet technologies have created a lot of pressure on the telecommunication industry. As a result telecommunication companies are working in highly challenging and competitive environment nowadays. Large amount of data is generated from various systems and this data is used for solving business problems that need urgent problem solving and handling. These data include call detail data, customer data and network data. Data Mining technology and business intelligence (BI) techniques are widely used for handling the business related problems. The key application areas of Data Mining in telecommunication industry are fraud detection, network fault isolation and improving market effectiveness. The aim of this paper is to explore various data mining tools and techniques and to check how they can be used to detect telecommunication fraud, fault and improve market effectiveness.**

Keywords: Data Mining, Telecommunications, Fraud Detection, Network Fault Isolation, Marketing & Churn

I. Introduction

The evolution of technology has enabled collection and storage of huge amounts of data. The size of databases today can be very large and can range up to terabytes. Practically it is impossible to analyze such large volumes of data using traditional techniques. For this reason data mining has gained a lot of attention. We can find and explore data, generate results, and learn from data. [1]

Data mining is the process of automatically finding and extracting useful information in large amount of data repositories. It is supported by:

1. Data availability.
2. Affordable processing power.
3. Inexpensive data storage.
4. Many commercial data mining tools are available.

Applying data mining to any industry depends on following two factors: the data that are available and the business problems faced by the industry. Data mining can be applied to any business, but in this paper we will describe how it can be used in the telecommunications industry. Telecommunications industry is data-dependent. Telecom companies have the detail records of the customers and the calls made by customers. [2]

The data mining process involves pre-processing of data, analyzing the data, post-processing of the data and the evaluation of results. For each of these steps some method is chosen based on the requirements and at the end of the process useful data is converted into information. This is depicted in the diagram below.

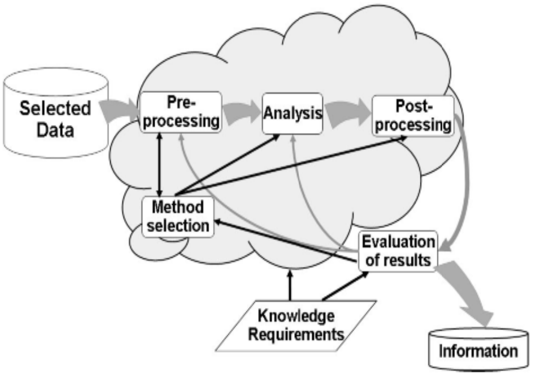


Fig. 1 Structure of Data Mining Process

A telecommunications network provides mobile services for a large geographical area. This area can cover, for example, a whole country. A mobile-phone user can move around the area without losing his connection to the network. This is achieved by placing Base Transceivers Stations (BTS) to give continuous coverage all over the area. A BTS has one or more transmitter-receiver pairs called Transceivers (TRX). Transceivers beam through BTS antennas, which are pointed to cover an area called a cell. When a mobile phone makes a call, it creates a connection to a BTS. From the BTS the call is typically forwarded to a transmission network, which connects all the BTSs to other network elements and to outside networks like Public Switched Telephone Networks (PSTN). When the match occurs in any of the BTS, the call is connected. [4]

This paper is organized as follows. First of all we describe how data mining can be used in telecommunications. Next we describe various types of telecommunication data. Next we explain various applications of data mining in telecommunications i.e. fraud detection, network isolation and marketing. Then we specify the work dome in this field. Next section specifies various issues and challenges and future trends. Final section gives a conclusion of the paper.

Telecommunications network management requires rapid decision-making, which can be supported by data mining methods. The decision-making is based on information extracted from large amounts of data that are continuously collected from networks. [4]

For a telecommunication organization to achieve dominance in market, following three strategies are open [7]:

1. They must excel at being the best low cost provider of services, making them the market dominator.
2. They be superior at quality of what they are doing, distinguishing them as best provider of the service.
3. They must become best marketers of service by pursuing "customer intimacy" i.e. being a company that responds to customer needs and wants better than anyone else.

II. II. Types of Telecommunication Data

The first step in the process of data mining is to understand the data. Without understanding, development of useful applications may not be possible. In this section we describe the three main types of telecommunication data. If the raw data is not suitable for data mining, then transformation of data is done to generate data that can be mined. [3]

A. *Call Detail Data*

Whenever a call is made on a network, the information about the call is saved as a call record. The call record contains information which describes the important features of every call. Such record usually consists of the originating and terminating phone numbers, the date and time of call and the duration of the call. The call record detail of every customer must be put into a single record that describes the customer's calling behaviour for extracting useful knowledge. This helps in generating customer profiles which can be used for mining data for marketing purposes. [5] Information that can be obtained from the call detail record include

1. Average call duration.
2. Average number of call received per day.
3. Average number of call originated per day.
4. Percentage of no-answer calls.
5. Percentage of weekday calls (Monday – Friday).
6. Percentage of day time calls (office hours)[6]

B. *Network Data*

A Telecommunication network consists of different types of equipment, which are made of many interconnected components. Each of these components can generate a status and error message that generates huge quantity of network data. The data is normally stored and analyzed so as to support network management functions such as fault isolation and detection. Data mining technology helps to perform these functions by automatically extracting knowledge from the network data. [5]

C. *Customer Data*

Telecommunication industries maintain a huge database of information of all their customers. The information consists of names, address, service plan, contract information, credit amount and payment history. These data are often used along with other data for example using customer data along with call detail data to identify phone frauds. Information about the customer can include [6]:

1. Name of the customer.
2. Address of the customer.
3. email id, additional contact nos.
4. Payment history.
5. Service plan etc.

III. Applications of Data Mining in Telecommunications

The main factors on which Data Mining applications depend include the details of the problem to be solved by the Data Mining and the availability of Data for mining. The main reason for the importance of Data Mining applications in the Telecommunications industry is the availability of large volumes of data. [6] There are many prospects of data mining techniques in telecommunication which may include predicting which customers are likely to default on payment, catching fraudulent activities, identifying of telecommunication patterns, improving resource utilization and service quality and allowing multi-dimensional data analysis to enhance the understanding of customer behaviour. Information obtained from data mining techniques can be used for application like market analysis, fraud detection, science exploration and retention of customer to production control. Data mining helps to reduce company's losses by making good predictions about the possible business outcomes. [5]

There are three major applications of data mining in telecommunication industry. These are:

1. Fraud Detection.
2. Network Fault Isolation & Prediction.
3. Marketing and customer relationship management (CRM).

A. *Fraud Detection*

Fraud is very serious issue faced by the telecommunication industry since it leads to the loss of revenue which may range to billions of dollars. As defined by Gosset & Hyland 1999, the telecommunication fraud can be stated as —any activity of using the telecommunication service intention of paying the company providing the service. [6] Fraud detection is important to the telecom industry because companies providing telecommunications services lose a significant amount of revenue due to frauds. [8] In order to identify fraud, data mining application can be used to analyze large amount of cellular call data which can be used to generate possible patterns. These patterns specify a customer's behaviour with respect to some pattern of fraud. The monitors are then fed into a neural network that determines when there is an evidence of fraud to raise an alert. Data mining also helps in detecting fraud by identifying and storing the phone numbers known to be used fraudulently. [5]

Telecommunication fraud can be classified into two main types Subscription fraud & Superimposition fraud. When a customer opens an account with the intention of never paying it is said to be Subscription fraud and when a unauthorized person gains illegal access to the account of a legitimate customer Superimposition fraud is said to occur. Telecommunication companies consider that superimposition frauds are the most significant problems. Both subscriptions fraud and Superimposition fraud should be detected as soon as possible and customer account should be deactivated immediately. [6]

Customer data can also be used for fraud detection. For example credit information and price plan can be in used in fraud analysis. Another method commonly for fraud detection is to create profile of customer's calling behaviour and compare activity against this behaviour. This behavior can be generated by defining the call detail records for a particular customer. Fraud can be detected immediately after it happens if the call details records for that customer are updated regularly. Fraud detection system work at customer level, not for individual call. Fraud detection involves predicting a rare event where the class distributions involved is complex. [6] Hence, data mining can be used to avoid loss revenues of telecommunication operator due to fraud.

B. *Data Mining Techniques for Fraud Detection*

The following data mining techniques can be used for fraud detection:

1) *Neural Networks*

Neural Networks calculate user profiles in an independent manner, thus adapting easily to the behaviour of the various users. [9] In order to differentiate between legitimate user and fraud, feed-forward neural networks (FF-NN) can be used. The problem involves the need to adapt the profiles so that the input-output mapping corresponds to the input-output pairs provided for that profile. The evaluation of performance is done by a Receiver Operating Characteristic (ROC) curve. ROC curve is graphical representation of the trade off between the true positive rates and the false positive rates for every possible cut off point separating overlapping distributions. [8]

2) *Decision Trees*

Learning algorithms commonly use divide-and-conquer approach. The input space is divided to maximize information gain or specify expression of knowledge change. This approach leads to tree-like data structures. The aim is to have leaves that contain objects of the same class. In particular, fraud cases are characterized by lower deviation values than normal use, which specifies that fraudsters show some kind of "compact" behaviour. They tend to place long calls. [8] For example consider a decision tree for the weekly representation of the users:

IF MeanCalls<0.86 THEN class=1 confidence: 71.98%, coverage: 70.48%)

IF MeanCalls>0.86 AND StdDur<129.5 THEN class=2 (conf.: 97.5%, cov.: 41.5%).

The first rule says that if a user places less than 1 call per day his is a legitimate user with confidence 72%. According to the second rule, if the mean number of calls in a week is more

that 1 (that is at least 7 costly calls in the week) and the standard deviation of their duration is less than 2 minutes, then the user is a fraudster with confidence 97.5%.

3) *Agglomerative Clustering*

No matter how well a neural network classifier may have performed, still there is no clue about the features that are actually used in order to achieve its performance. In order to further investigate the problem of appropriate user modeling for fraud detection, the hierarchical agglomerative clustering technique can be applied on the data. The aim is to test whether cases from the same class tend to form clusters and if yes, then under which condition. During hierarchical agglomerative clustering the user does not specify the expected number of clusters *k*. Instead, the algorithm constructs a tree-like hierarchy, a dendrogram, which implicitly contains all values of k. The root of the tree structure defines a cluster that contains all data, while its leafs represent the *n* clusters, each one containing one of the *n* objects. The agglomerative clustering algorithm starts with each object representing a cluster, called a singleton, and proceeds by fusing the closest ones until a single cluster is obtained. Therefore, a measure of dissimilarity between two clusters must be defined. Two different distance measures, namely the Euclidean distance and the correlation between objects, are used. Clustering quality can be judged by means of appropriate statistics such as the agglomerative coefficient (AC) and the cophenetic coefficient (CC). [8]

C. *Network Fault Isolation & Prediction*

Telecommunication networks consist of complex configurations of software and hardware. Since the industry requires network efficiency and reliability, most of the network elements are designed to generate status and alarm messages in case of any problem. Expert systems were designed to handle such alarms. Network fault isolation in the telecommunication industry is a quiet difficult task because of the following reasons. Huge volumes of data are available and a single fault can generate different alarms which may not be related with each other. Hence alarm correlation is very important in predicting network faults. An active and rapid response is very much essential for maintaining the reliability of the network. Data mining techniques like neural networks, classification and sequence analysis can be used for identifying faults. The telecommunication Alarm Sequence Analysis (TASA) is a Data Mining tool which provides fault identification by identifying recurrent patterns in algorithms. It can be used to generate an alarm correlation system, which can be used to identify faults. Genetic algorithms can also be used to predict telecommunication failures. Time weaver is such a genetic algorithm which has the capability to operate directly on the raw network data. Standard tools of classification can be used to predict future failures but it has drawbacks like, some of the data may be lost in process. [8]

D. *Marketing*

Telecommunication industries maintain huge amount of information about their customers. Hence, they can use data mining to identify and retain customers and maximize the profit from each customer. [7] Data mining also helps in designing customer's profiles from call record details and then mining these profiles for marketing. The emphasis of marketing in telecom has moved from identifying new customers to measuring customer value and then taking necessary steps to retain profitable customers. [5] Numerous Data Mining techniques can be used to generate customer life time value. To estimate the life time value of a customer we need to estimate how long they will remain with current network. It will help in predicting when a customer is likely to leave and then taking necessary steps to retain the customer. One serious issue that the telecom industries face is customer churn.

Predicting churn, i.e. to check if a customer may leave for a competitor, is an important application of analyzing customer behavior. It is expensive to get new customers then to retain existing ones. Correctly predicting if a customer is about to churn and then convincing him to stay can increase the revenue of a company. [10] Customer churn is used to specify the movement of customer from one provider to another, and 'churn management' describes an operator's process to retain exiting profitable customers. [12]

Telecommunication industry is a large market and is aware of the importance of Customer Management and Relationship and the impact of Churn. Because of this, developments on following fields are achieved: [11]

Cross Selling and up-selling: maximize profits from existing customers.

*Retaining and up-selling***:** retaining profitable customers or get rid of inappropriate customers in the company profile.

Poaching: to poach (get) new Customers from rival companies.

Obtaining new customers is relatively expensive than retaining existing customers. It is for this reason that telecommunication companies realize that keeping existing customers is more important and churn analysis is an important data mining application areas.

A Person who Churns from one network to another is called a Churner. Two categories of churners are voluntary and involuntary churners: These are explained below [13]:

1. *Involuntary churners* are customers that the company decides to remove from its list of subscribers. This category includes people that are removed for non-payment (customers with credit problem), frauds (customers who cheat), and under-utilization (customers who don't use the phone).
2. *Voluntary churn* occurs when the customer terminates the service. We recognize two types of voluntary churn: incidental and deliberate churn.

Incidental churn occurs, not because something happened in customers lives and not because the customers planned on it. For example: change in financial condition churn, change in location churn, etc.

Deliberate churn occurs due to reasons of technology (customers wanting newer and better technology), economics (price and charges), service quality, social or psychological factors, and convenience reasons. Deliberate churn is the problem that most companies try to solve.

Post paid and Prepaid Churn When considering Post paid churn, the deactivation date, i.e. the date when the customer is disconnected from the network, is equal to the churn date. Because, this is the date when a customer stops using the operator's services. Whereas prepaid segment does not have contract between users and telecom operator, so the definition of Prepaid churn is not simple. However in Prepaid churn, the deactivation date may not match the churn date. [13] Generally it takes a long period of time before a Prepaid customer is disconnected from network. In many cases customers may have churned long before they are disconnected from network. This is the reason why the deactivation date is not suitable indicator for prepaid churn.

Following steps can be followed for the churn prediction [14]:

1. Initially, for each attribute, a threshold value is assigned.

2. The attribute values of the training dataset are compared with the attribute's threshold to declare that a customer will churn or not. Simple if…then …else rules are applied in this process.
3. A model is then constructed for the training dataset.
4. The model is then applied on the test dataset and the results are listed.
5. The above steps can be repeated by varying the threshold values of the attributes selected.

Churn prediction data mining assessment methodology
The purpose of this research is to assess the performance of various data mining techniques when used in context of churn prediction. The methodology consists of three parts [12]:

1. An IT infrastructure, which includes a common customer base, attributes and transactions, modeling parameters, model results, etc.
2. A model-independent knowledge discovery procedure to discover customer behaviour prior to churn, by using data mining techniques.
3. A set of measurements to measure the performance of models developed by different modeling tools, such as decision tree and neural network.

IV. Related Work

In 2006, Shin-Yuan Hung [12] studied how data mining can be applied to telecommunication for churner predication and what measures need to be undertaken to ensure that churn does not occurs.

Sen Wu, Naidong Kang [2] in 2007 stated that customer data available should be used by the operator to analyze the common characteristics of fraudulent behaviour of customers in telecom industry systematically for detection frauds.

D. Camilovic [1] in February 2008 emphasised that CRM and churn cannot be practiced in business without tracking patterns within customer data. It highlights the importance of using the data available to predict churn.

Gary M. Weiss [3] in 2009 explained various types of telecom data and how data mining can be used in applications like marketing (for CRM and Churn), fraud detection and network isolation and for future predictions.

Rob Mattison [7] in his book explains what data mining is and how it can be used for marketing in telecommunications for Churn management.

Kimmo H¨at¨onen [4] in 2009 in his thesis explained how decision making can be implemented on various levels for making decisions and identifies the data mining tools that can be used for this purpose.

Umman Tugba Simsek Gürsoy [11] in 2010 provides an analysis of churn and its impact on the operator and explains what techniques should be used to predict and prevent it.

Frank Eichinger [10] explains what churn is and how we can predict churn in a network and what techniques to use to avoid churn.

Goran Kraljevi´c [13] in 2010 explained how prepaid churn analysis is difficult as compared to postpaid churn and how to deal with prepaid churn using data mining tools and techniques.

Isinkaye O. Folasade [5] in 2011 explores different data mining tools and applications and how they can be used to detect telecommunication fraud, fault and improve market effectiveness.

Constantinos S. Hilas [8] in March 2012 explains how data mining can be applied over data available to an operator for checking abnormal data patterns which may be helpful in detecting fraudulent customers and data.

V. Umayaparvathi [14] in March 2012 explained in her journal what techniques to use for predicting churn in telecommunication environment and what techniques to use to ensure that a customer does not leaves the operator.

Madhuri V. Joseph [6] in February 2013 explained that telecommunication industries have large amount of data(call detail data, customer data and network data) which can be used for solving many business problems that require urgent handling.

Anita B. Desai [9] in 2013 in her journal explains how data mining techniques like neural networks, decision trees and agglomerative clustering can be used for detecting frauds.

V. Future Trends and Challenges

Data Mining has an important role because the data is the primary concern in telecommunication industry. The telecommunication data is mainly in the form of transactions or events and is not at a level for mining so complex pre-processing of data is required. Scalability is an important issue because the industry is handles very large databases. Fraud detection and network fault isolation in the telecom increase the importance of real-time operation. Data Mining should consider the privacy issues. This is so because telecommunication companies maintain private information like whom each customer calls. Another issue with telecom data and its applications involves rarity. Predicting and identifying rare events like fraud and equipment failure is quite difficult for many data mining algorithms and therefore this issue must be handled carefully in order to ensure good results. It is certain that the new Data Mining applications will be developed and deployed which will help to eliminate some of the problems faced by current applications.

VI. Conclusion

Data mining applications play an important role in the telecom industry due to the presence of large amount of data and the rigorous competition faced in the sector. The primary application areas are marketing and Customer Relationship Management, Fraud detection, Network Management. The latest developments in the Data Mining involve enhancement and implementation of existing techniques and methods to ensure continuous growth of telecommunication companies that using them. This paper explains how data mining tools and techniques can be used by telecommunication companies to discover and extract useful patterns from large volumes of data so as to find observable patterns, which can help in identifying and catching fraudulent activities, improving resource utilization and service quality, facilitating multi-dimensional data analysis to improve the understanding of customer behaviour.

ACKNOWLEDGMENT

We would like to thank almighty for his constant blessings. Then we would like to dedicate our gratitude towards parents, teachers, family, friends, and in essence, all sentient beings.

REFERENCES

[1] D. Camilovic, “Data Mining And CRM In Telecommunications” Serbian Journal of Management 3 (1), February 2008.

[2] Sen Wu, Naidong Kang, Liu “Yang Fraudulent Behavior Forecast In Telecom Industry Based On Data Mining Technology” Communications of the IIMA 1, Volume 7 Issue 4 (3), 2007.

[3] Gary M. Weiss, “Data Mining In The Telecommunications Industry”, IGI Global, 2009.

[4] Kimmo H¨at¨onen, “Data Mining For Telecommunications Network Log Analysis”, Helsinki University Printing House, 2009.

[5] Isinkaye O. Folasade, “Computational Intelligence In Data Mining And Prospects In Telecommunication Industry”, Journal of Emerging Trends in Engineering and Applied Sciences (JETEAS) 2 (4) (ISSN: 2141-7016), 2011.

[6] Madhuri V. Joseph, “Data Mining and Business Intelligence Applications In Telecommunication Industry”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-3, February 2013.

[7] Rob Mattison, “Data Warehousing And Data Mining For Telecommunications”, ISBN 0-89006-952-2, Artech House Boston London.

[8] Constantinos S. Hilas, “Data Mining Approaches To Fraud Detection In Telecommunications”, 2nd Pan-Hellenic Conference on Electronics and Telecommunications PACET’12, March 2012.

[9] Anita B. Desai, Dr. Ravindra Deshmukh “Data Mining Techniques For Fraud Detection”, International Journal of Computer Science and Information Technology (IJCSIT) ISSN: 0975-9646, Volume 4(1), 2013.

[10] Frank Eichinger, Detlef D. Nauck, and Frank Klawonn, “Sequence Mining For Customer Behaviour Predictions In Telecommunications”.

[11] Umman Tugba Simsek Gürsoy, (2010) “Customer Churn Analysis In Telecommunication Sector”, Istanbul University Journal of the School of Business Administration, ISSN: 1303-1732, Vol:39, No:1[12] Shin-Yuan Hung, David C. Yen, Hsiu-Yu Wang, “Applying Data Mining To Telecom Churn Management”, Expert Systems with Applications, 2006.

[12] Goran Kraljevi´c, Sven Gotovac, “Modeling Data Mining Applications For Prediction Of Prepaid Churn In Telecommunication Services”, AUTOMATIKA 51, ISSN 0005-1144, 2010.

[13] V. Umayaparvathi, K. Iyakutti, “Applications Of Data Mining Techniques In Telecom Churn Prediction”, International Journal of Computer Applications (0975 – 8887) Volume 42– No.20, March 2012.

Network Programmability Using POX Controller

Sukhveer Kaur[1], Japinder Singh[2] and Navtej Singh Ghumman[3]
[1,2,3]Department of Computer Science and Engineering,
SBS State Technical Campus, Ferozepur, India
E-mail: [1]bhullarsukh96@gmail.com, [2]japitaneja@gmail.com,
[3]navtejghumman@yahoo.com

***Abstract*—POX is a Python based open source OpenFlow/Software Defined Networking (SDN) Controller. POX is used for faster development and prototyping of new network applications. POX controller comes pre installed with the mininet virtual machine. Using POX controller you can turn dumb openflow devices into hub, switch, load balancer, firewall devices. The POX controller allows easy way to run OpenFlow/SDN experiments. POX can be passed different parameters according to real or experimental topologies, thus allowing you to run experiments on real hardware, testbeds or in mininet emulator. In this paper, first section will contain introduction about POX, OpenFlow and SDN, then discussion about relationship between POX and Mininet. Final Sections will be regarding creating and verifying behavior of network applications in POX.**

Keywords: POX, SDN, Open Flow, Mininet

I. Introduction

SDN separates the control plane of networking device (switch/ router) from its data plane, making it possible to control, monitor, and manage a network from a centralized controller.

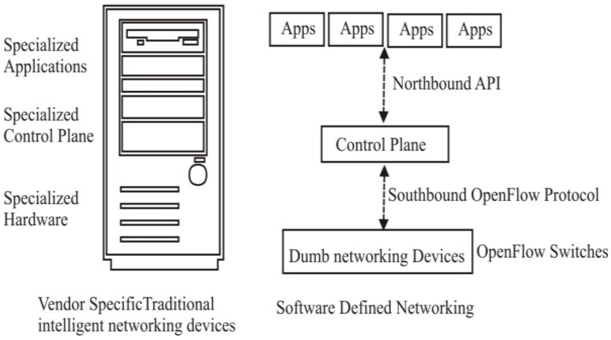

Specialized Applications
Specialized Control Plane
Specialized Hardware
Vendor SpecificTraditional intelligent networking devices
Apps
Apps
Apps
Apps
Northbound API
Control Plane
Southbound OpenFlow Protocol
Dumb networking Devices
OpenFlow Switches
Software Defined Networking

Fig. 1 Decoupled Control and Data Plane

Software Defined Networking [1] tries to simplify the development of new applications by separating the data plane from control plane. Control plane is also called controller. This controller has a global view of the network and controls the flow through the network. Since most intelligence is now transferred to the controller, the switch only perform the actions that the controller requests. This makes the switches very simple and inexpensive. But in traditional networks (Fig. 1), each device has vendor-specific operating system to control the data plane. Additional applications can be implemented on top of this operating system.

POX [2] is an open source controller for developing SDN applications. POX controller provides an efficient way to implement the OpenFlow protocol which is the de facto communication protocol between the controllers and the switches. Using POX controller you can run different applications like hub, switch, load balancer, and firewall. Tcpdump packet capture tool can be used to capture and see the packets flowing between POX controller and OpenFlow devices.

Communication between the controller and the switches is carried by communication protocol such as OpenFlow [3], ForCES [4] (Fig. 2). OpenFlow is the most popular standard protocol used in SDN. OpenFlow switches behave as dumb forwarding devices. They are unable to perform any actions without programmed by the controller.

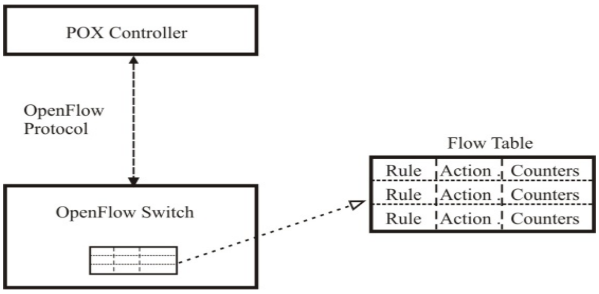


Fig. 2 POX Controller

When a switch is powered on, it will immediately connect to an OpenFlow controller. Initially, the flow table of the switches is empty. When a packet arrives at a switch, it does not know, how this packet is to be handled. Then it send packet-in message to the controller. To handle the packet, controller inserts a flow entries in flow table of switch. Flow entry in flow table contains three parts, rule(match field), action, counters. For each packet, that has to pass through a switch, a flow entry will have to be installed so that the switch can forward this traffic without further intervention of the controller [5]. Flow modification messages are sent to the switches to install the flow entries in flow table (Fig. 3). Once these are installed, traffic belonging to this flow will be handled by the switches themselves.

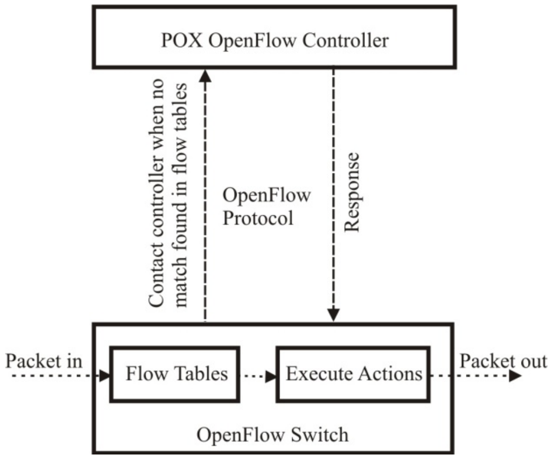


Fig. 3 SDN Architecture

II. POX AND MININET

Mininet is an emulation tool that allows running a number of virtual hosts, controllers, switches, and links. It uses container based virtualization to make a single system act as a complete network. It is a simple, robust and inexpensive network tool to develop and test OpenFlow based applications. Mininet [6] can create a complex network topology for testing purposes, without configuring the physical networks. It supports custom topologies. It supports simple and extensible Python API for network creation and testing.

Mininet combines the desirable features of simulators, testbeds and emulators. Mininet is cheaper, easily available, and quickly reconfigurable as compared to testbeds such as GENI [7], VINI [8], and Emulab [9]. It runs real, unmodified code as compared to simulators such as EstiNet [10], ns-3 [11]. The code that is to be developed in Mininet, can also run in real network without any modifications. It supports large scale networks containing large number of virtual hosts and switches. In short, Mininet's virtual hosts, switches, links, and controllers are just like the real thing. They are just created using software rather than hardware.

Mininet have built-in Controller classes to support different network controllers such as reference controller (controller), ovs-controller [12] and less used NOX Classic [13].

You can choose controller by invoking 'mn' command.

```
# mn --controller ref
# mn --controller ovsc
# mn --controller nox
```

Five most important open source controllers (Table I) that can be used by Mininet remotly are POX, Ryu [14], Trema [15], FloodLight [16], and OpenDaylight [17]. There are number of other SDN controllers like NOX (C++) , Jaxon (Java) [18], Beacon (Java) [19], Maestro (Java) [20] which are not considered because they are deprecated and poorly documented.

A. POX

The POX is a python based SDN controller that is inherited from the NOX controller.

B. Ryu

Ryu is a component-based SDN controller. Ryu has a collection of built-in components. These components can be changed, extended and composed for creating new customized controller applications. Any programming language can be used to develop a new component.

C. Trema

Trema is a framework for Ruby and C that builds software platform for OpenFlow developers. It is easy to use Open Source free software.

TABLE 1 DIFFERENT SDN CONTROLLERS

	POX	Ryu	Trema	Floodlight	Open Day Light
Language Support	Python	Python	C Ruby	Java	Java
OpenFlow Support	v1.0	v1.0 v1.2 v1.3	v1.0	v1.0	v1.0
OpenSource	Yes	Yes	Yes	Yes	Yes
GUI	Yes	Yes	No	Web GUI	Yes
REST API	No	Yes	No	Yes	Yes
Platform Support	Linux Mac Windows	Linux	Linux	Linux	Linux Mac Windows

D. Floodlight

The Floodlight Open SDN Controller is an Apache licensed, enterprise class, Java based OpenFlow controller. FloodLight controller contains a number of modules, where each module provides a service to the other modules and to the control logic application through simple Java API or a REST API.

E. OpenDayLight

OpenDayLight is an open source project. The goal of the project is to create robust code that covers major components of the SDN architecture, to gain acceptance among the vendors and users, and to have a growing community that contributes to the code and uses the code for commercial products.

To use POX controller, type the following command in terminal window.

```
# python pox.py log.level –DEBUG
```

Using this command POX controller runs in DEBUG mode. DEBUG mode allows display of additional messages exchanged with the switch. To launch Mininet with default topology of 1 switch and 2 hosts run the following command.

```
# mn
```

In this case switch will connect to the default ovs controller. If you want to use POX controller running on the same Mininet machine you need to run the 'mn' command with the controller option having the parameters set to 'remote'. Loopback address '127.0.0.1' will be used as ip address. The following command will connect the switches to remote POX controller running on another terminal.

```
# mn –controller=remote, ip=127.0.0.1
```

But if POX controller is on different machine (suppose 172.24.0.1), then run the following command

```
# mn --controller=remote, ip=172.24.0.1
```

You can create complex pre defined or custom defined topologies using Mininet. For example

```
# mn –mac --topo single,5 --switch ovsk --controller remote
```

This will create 5 hosts and 1 switch topology. The different options that can be used with 'mn' command are shown in Table II.

TABLE 2 MININET OPTIONS

Commands	Description
mn	run Mininet
--topo single, 5	create 1 switch with 5 hosts
--mac	makes mac address same as node number on hosts
--arp	install static ARP entries
--switch ovsk	use Open vSwitch
--controller remote	use remote controller
--ip	remote controller ip address

III. POX APPLICATIONS

There are various applications that can be created using POX. The network application could be a simple hub, switch, and router or could be sophisticated middle boxes such as firewall or load balancer. This section contains simple hub logic and application code.

A. *Hub Application*

If a flow entry in flow table contains action to flood the packet that arrives at specific port of forwarding device, then that device act like a hub. In the topology shown in Fig. 4, all hosts belong to the same network. When host h1 wants to send a packet to host h4, then it first sends a packet to forwarding device at port 1. When a packet arrives at port 1, then it matched against flow entry. When match is found, then it is flooded to all ports except the incoming port according to action specified in flow entry. If no match is found, then packet is forwarded to controller. In one terminal window, run the following command to create an experiment topology.

```
# mn --mac --topo single,5 --switch ovsk --controller remote
```

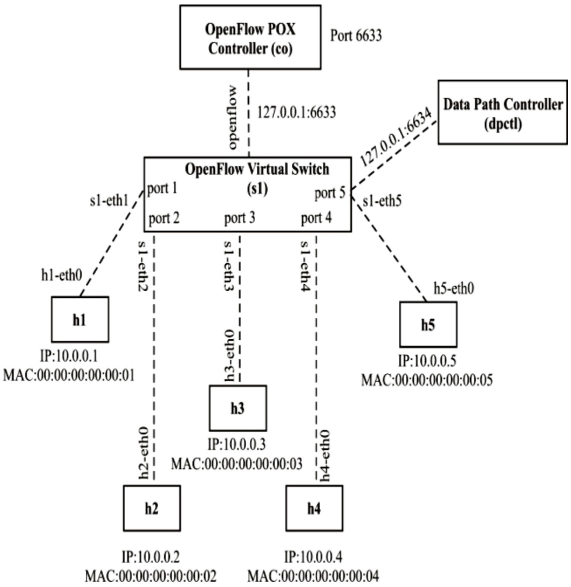
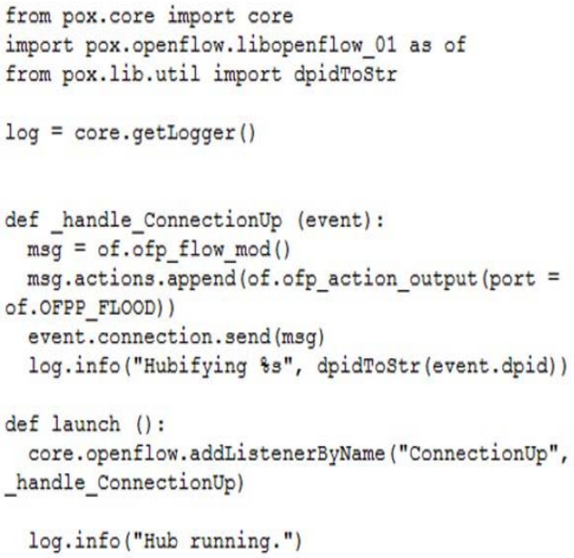


Fig. 4 Single Switch, 5 Hosts Topology



```
from pox.core import core
import pox.openflow.libopenflow_01 as of
from pox.lib.util import dpidToStr

log = core.getLogger()


def _handle_ConnectionUp (event):
  msg = of.ofp_flow_mod()
  msg.actions.append(of.ofp_action_output(port =
of.OFPP_FLOOD))
  event.connection.send(msg)
  log.info("Hubifying %s", dpidToStr(event.dpid))

def launch ():
  core.openflow.addListenerByName("ConnectionUp",
_handle_ConnectionUp)

  log.info("Hub running.")
```

Listing 1 Hub Application Code

This will launch Mininet network topology consisting of 1 OpenFlow switch, 1 OpenFlow controller and 5 hosts. The POX controller comes pre-installed with the provided VM image. From another terminal window, run the hub code (code file name is 'hub.py') shown in Listing 1 by using the following command.

```
# python pox.py log.level --DEBUG hub
```

This will launch the POX controller in verbose mode for debugging purposes and also run the hub application.

B. *Understanding Hub Application Code*

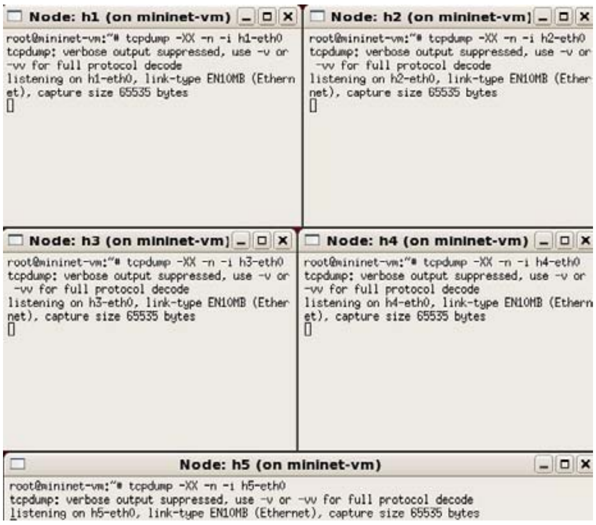
Before implement a hub application, first you need to import a core object that show a connection between modules in POX and OpenFlow library that is used for access a number of primitives.

1. ofp_action_output class: This class specifies a switch port, where you want to send the packet. There are various "special" port numbers. For example in the hub application, ‘OFPP_FLOOD’ which sends the packet to all ports except the incoming port.
2. ofp_flow_mod OpenFlow message: This message is send from controller to switch to insert flow table entry. Flow table entries will be matched against fields of incoming packets and then perform some actions on matching packets.
3. connection.send(...): Controller sends an OpenFlow message to a switch by using this function. A ‘ConnectionUp’ event is fired, When a connection to a switch starts. The above code call a ‘_handle_ConnectionUp ()’ function that contains hub logic.
4. launch(): The launch() function is automatically called, when the application is started. The application registers all event listeners in this function.
5. dpid_to_str(): Each OpenFlow switch has a unique 64 bit datapath ID (DPID) and that is to be passed to controller from switch during handshaking. 48 bits are Ethernet address and 16 bits are implementation defined. It is a decimal number that is not easy to understand. POX define a pox.lib.util.dpid_to_str () function to format DPIDs.

IV. VERIFYING HUB BEHAVIOR

To verify hub behavior, Start a topology that contains single switch and 5 hosts and run it with POX controller. From host h1 sends icmp packets to the host h3. Here all the hosts see the same exact traffic which is the default behavior of hub. Launch an ‘xterm’ for each host and view the traffic simultaneously for each host by using ‘tcpdump’ (Fig. 5). For this purpose start 5 xterm terminals, one for each host.

The command for viewing traffic is “tcpdump”. Pass the option ‘-XX’ for verbose output, ‘-i’ for specifying the interface for listening, ‘-n’ for no name resolution.



```
root@mininet-vm:~# tcpdump -XX -n -i h1-eth0
tcpdump: verbose output suppressed, use -v or
-vv for full protocol decode
listening on h1-eth0, link-type EN10MB (Ethern
et), capture size 65535 bytes
```

```
root@mininet-vm:~# tcpdump -XX -n -i h2-eth0
tcpdump: verbose output suppressed, use -v or
 -vv for full protocol decode
listening on h2-eth0, link-type EN10MB (Ether
net), capture size 65535 bytes
```

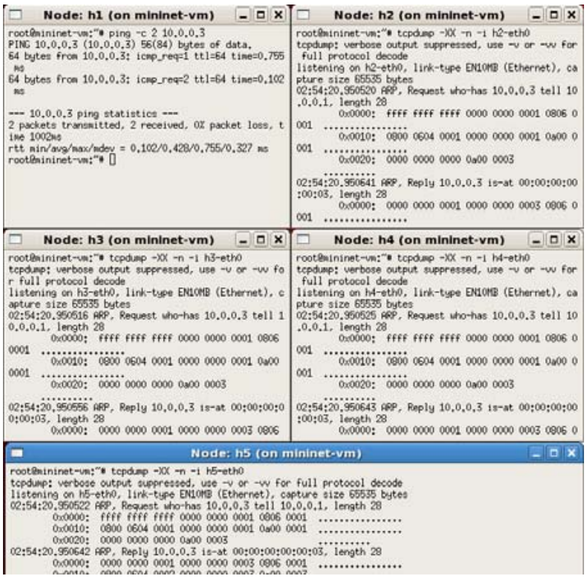
```
root@mininet-vm:~# tcpdump -XX -n -i h3-eth0
tcpdump: verbose output suppressed, use -v or
 -vv for full protocol decode
listening on h3-eth0, link-type EN10MB (Ether
net), capture size 65535 bytes
```

```
root@mininet-vm:~# tcpdump -XX -n -i h4-eth0
tcpdump: verbose output suppressed, use -v or
-vv for full protocol decode
listening on h4-eth0, link-type EN10MB (Ethern
et), capture size 65535 bytes
```

```
root@mininet-vm:~# tcpdump -XX -n -i h5-eth0
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on h5-eth0, link-type EN10MB (Ethernet), capture size 65535 bytes
```

Fig. 5 Tcpdump Outputbefore Running 'Ping' Utility

Now from host h1 ping to the host h3 at address 10.0.0.3. Ping packets will first go to the controller, which will then flood the packets to all hosts except the interface which sent the packet. You will see identical ICMP and ARP packets related to the ping in all the terminals. (Fig. 6) thus verifying the behavior of hub.



```
root@mininet-vm:~# ping -c 2 10.0.0.3
PING 10.0.0.3 (10.0.0.3) 56(84) bytes of data.
64 bytes from 10.0.0.3: icmp_req=1 ttl=64 time=0.755
 ms
64 bytes from 10.0.0.3: icmp_req=2 ttl=64 time=0.102
 ms

--- 10.0.0.3 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, t
ime 1002ms
rtt min/avg/max/mdev = 0.102/0.428/0.755/0.327 ms
root@mininet-vm:~# 
```

```
root@mininet-vm:~# tcpdump -XX -n -i h2-eth0
tcpdump: verbose output suppressed, use -v or -vv for
 full protocol decode
listening on h2-eth0, link-type EN10MB (Ethernet), ca
pture size 65535 bytes
02:54:20.950520 ARP, Request who-has 10.0.0.3 tell 10
.0.0.1, length 28
        0x0000:  ffff ffff ffff 0000 0000 0001 0806 0
001  ................
        0x0010:  0800 0604 0001 0000 0000 0001 0a00 0
001  ................
        0x0020:  0000 0000 0000 0a00 0003
     ..........
02:54:20.950641 ARP, Reply 10.0.0.3 is-at 00:00:00:00
:00:03, length 28
        0x0000:  0000 0000 0001 0000 0000 0003 0806 0
001  ................
```

```
root@mininet-vm:~# tcpdump -XX -n -i h3-eth0
tcpdump: verbose output suppressed, use -v or -vv fo
r full protocol decode
listening on h3-eth0, link-type EN10MB (Ethernet), c
apture size 65535 bytes
02:54:20.950516 ARP, Request who-has 10.0.0.3 tell 1
0.0.0.1, length 28
        0x0000:  ffff ffff ffff 0000 0000 0001 0806
0001  ................
        0x0010:  0800 0604 0001 0000 0000 0001 0a00
0001  ................
        0x0020:  0000 0000 0000 0a00 0003
      ..........
02:54:20.950556 ARP, Reply 10.0.0.3 is-at 00:00:00:0
0:00:03, length 28
        0x0000:  0000 0000 0001 0000 0000 0003 0806
```

```
root@mininet-vm:~# tcpdump -XX -n -i h4-eth0
tcpdump: verbose output suppressed, use -v or -vv for
 full protocol decode
listening on h4-eth0, link-type EN10MB (Ethernet), ca
pture size 65535 bytes
02:54:20.950525 ARP, Request who-has 10.0.0.3 tell 10
.0.0.1, length 28
        0x0000:  ffff ffff ffff 0000 0000 0001 0806 0
001  ................
        0x0010:  0800 0604 0001 0000 0000 0001 0a00 0
001  ................
        0x0020:  0000 0000 0000 0a00 0003
     ..........
02:54:20.950643 ARP, Reply 10.0.0.3 is-at 00:00:00:00
:00:03, length 28
        0x0000:  0000 0000 0001 0000 0000 0003 0806 0
```

```
root@mininet-vm:~# tcpdump -XX -n -i h5-eth0
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on h5-eth0, link-type EN10MB (Ethernet), capture size 65535 bytes
02:54:20.950522 ARP, Request who-has 10.0.0.3 tell 10.0.0.1, length 28
        0x0000:  ffff ffff ffff 0000 0000 0001 0806 0001  ................
        0x0010:  0800 0604 0001 0000 0000 0001 0a00 0001  ................
        0x0020:  0000 0000 0000 0a00 0003                 ..........
02:54:20.950642 ARP, Reply 10.0.0.3 is-at 00:00:00:00:00:03, length 28
        0x0000:  0000 0000 0001 0000 0000 0003 0806 0001  ................
```

Fig. 6 Tcpdump Output after Running 'Ping' Utility

V. CONCLUSION

POX controller can be used to convert cheap, dumb merchant silicon devices into hub, switch, router or middleboxes such as firewall, load balancer. POX is also great tool for deploying and testing SDN applications. Its great strength lies in that it can be used with real hardware, in testbeds or with Mininet emulator. The POX controller has some great features but does not have GUI interface. Open Flow v1.0 is most widely used version. Open Flow version 1.3 will be the next version that is supposed to be widely implemented in products. POX supports only v1.0. So support for v1.3 could be future challenge area. The network applications created in POX controller can not be used with other controllers. Porting of POX network applications to other controllers can be another research area.

ACKNOWLEDGMENT

The authors would like to thank Mr. Vipin Gupta from U-Net Solutions, Moga, India for his valuable assistance for this work.

REFERENCES

[1] Nunes, B.; Mendonca, M.; Nguyen, X.; Obraczka, K.; Turletti, T., "A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks," *Communications Surveys & Tutorials, IEEE* , vol.PP, no.99, pp.1,18.

[2] Fernandez, Marcial. "Evaluating OpenFlow controller paradigms." In *ICN 2013, The Twelfth International Conference on Networks*, pp. 151-157. 2013.

[3] Lara, Adrian, Anisha Kolasani, and Byrav Ramamurthy. "Network innovation using openflow: A survey." (2013): 1-20.

[4] Zhou, Lei, Ligang Dong, and Rong Jin. "Research on ForCES Configuration Management Based on NETCONF." *Information Technology Journal* 13, no. 5 (2014).

[5] Kim, Hyojoon, and Nick Feamster. "Improving network management with software defined networking." *Communications Magazine, IEEE* 51, no. 2 (2013): 114-119.

[6] Lantz, Bob, Brandon Heller, and Nick McKeown. "A network in a laptop: rapid prototyping for software-defined networks." In Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, p. 19. ACM, 2010.

[7] GENI at http://www.geni.net/

[8] VINI at http://www.fp7-federica.eu/pres_eventi/20081014-vini-bavier.pdf.

[9] M. Hibler, R. Ricci, L. Stoller, J. Duerig, S. Guruprasad, T. Stack, K. Webb, and J. Lepreau. Large-scale virtualization in the emulab network testbed. In USENIX 2008 Annual Technical Conference, pages 113-128. USENIX, 2008.

[10] Wang, Shie-Yuan, Chih-Liang Chou, and Chun-Ming Yang. "OpenFlow Controllers over EstiNet Network Simulator and Emulator: Functional Validation and Performance Evaluation."

[11] Henderson, Thomas R., Mathieu Lacage, George F. Riley, C. Dowell, and J. B. Kopena. "Network simulations with the ns-3 simulator." SIGCOMM demonstration (2008).

[12] OVS controller at http://yuba.stanford.edu/~casado/of-sw.html.

[13] N. Gude, T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown,and S. Shenker. Nox: towards an operating system for networks. ACM SIGCOMM Computer Commun. Review, 38(3):105–110, 2008.

[14] Shalimov, Alexander, Dmitry Zuikov, Daria Zimarina, Vasily Pashkov, and Ruslan Smeliansky. "Advanced study of SDN/OpenFlow controllers." In *Proceedings of the 9th Central & Eastern European Software Engineering Conference in Russia*, p. 1. ACM, 2013.

[15] Trema at https://github.com/trema/trema

[16] Kim, Hyojoon, and Nick Feamster. "Improving network management with software defined networking." *Communications Magazine, IEEE* 51, no. 2 (2013): 114-119.

[17] OpenDayLight at http://www.opendaylight.org/

[18] “Jaxon,” accessed 11-June-2013 at http://jaxon.onuos.org/

[19] Erickson, David. "The beacon openflow controller." In Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking, pp. 13-18. ACM, 2013.

[20] EugeneNg, ZhengCai AlanL Cox TS. "Maestro: Balancing Fairness, Latency and Throughput in the OpenFlow Control Plane."

Mininet as Software Defined Networking Testing Platform

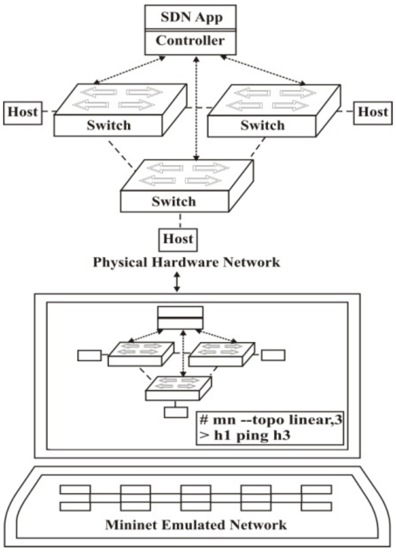
Karamjeet Kaur[1], Japinder Singh[2] and Navtej Singh Ghumman[3]
[1,2,3]Department of Computer Science and Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, India
E-mail: [1]bhullar1991@gmail.com, [2]japitaneja@gmail.com, [3]navtejghumman@yahoo.com

***Abstract*—Mininet is an emulator for deploying large networks on the limited resources of a simple single Computer or Virtual Machine. Mininet has been created for enabling research in Software Defined Networking (SDN) and OpenFlow. Mininet emulator allows running unmodified code interactively on virtual hardware on a simple PC. It provides convenience and realism at very low cost. The alternative to Mininet is hardware test beds which are fast, accurate but very expensive and shared. The other option is to use simulator which is very cheap but sometimes slow and require code modification. Mininet offers ease of use, performance accuracy and scalability.**

Keywords: Mininet, SDN, OpenFlow

I. Introduction

There is need to model hosts, switches, links and SDN/OpenFlow controllers. Mininet [1] allows creating topologies of very large scale size up to thousands of nodes and perform test on them very easily. It has very simple command line tools and API. Mininet allows the user to easily create, customize, share and test SDN networks. (Fig. 1).



```
# mn --topo linear,3
> h1 ping h3
```

Fig. 1 Emulating Real Networks in Mininet

Mininet is freely available open source software that emulates OpenFlow devices and SDN controllers. Mininet can simulates SDN networks, can run a controller for experiments [2]. It allows emulating real world network scenarios Couple of SDN controllers are included with in Mininet VM. The default controllers are good but for implementing advance concepts, POX [3] controller is used.

II. SDN and Open Flow

In traditional networks [4], the data plane and the control plane are tightly coupled on the same device (Fig. 2). Therefore in traditional networks, development of new applications and modification in behavior of existing devices is very difficult task. Software Defined networking (SDN) [5] overcomes these problems by shifting the control logic from devices to the centralized place. The shifted control logic is called SDN Controller or Network Operating System (NOS) [6]. The controller has a global view of the entire network, Therefore by using SDN you can manage the functionality of network in a very efficient manner.

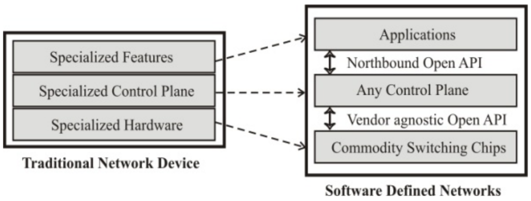


Fig. 2 Separate Control and Data Plane

Open Flow [7] is a standard protocol that is used to provide a communication between controller and dumb device. The controller and dumb devices are called control plane and data plane respectively. The Open Flow controller is responsible for deciding which action is to be performed by the switch. The decision approach is either Reactive or Proactive.

In the Reactive approach, when a packet arrives at a switch, switch does not know how to handle that packet. Therefore switch sends the packet to the controller [8]. Controller is responsible for inserting a flow entry into the flow table of a switch using the openflow protocol. The main disadvantage of this approach is that switch is totally dependent upon controller decision. So when a switch loses the connection with the controller, it cannot handle that packet.

In the Proactive approach [9], the controller pre populates the flow entries in the flow tables of each switch. This approach overcomes the limitation of reactive approach because even if the switch loses the connection with controller, it does not disrupt traffic.

The main advantages of SDN over traditional approach are that it allows you to quickly test and deploy new applications in real network, minimize capital and operating expenses and allows centralized management of each switch.

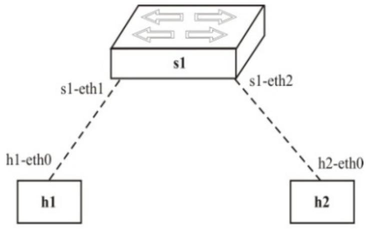
III. Mininet Topologies

Mininet contains number of default topologies such as minimal, single, reversed, linear and tree [10]. This section explains these topologies one by one. Understanding naming method for interfaces, hosts and switches is essential for prospering using Mininet. Switches are named from s1 to sN. Hosts are named h1 to hN. Host interfaces are named prefixed with host's name following by Ethernet name starting with 0. First interface of host 'h1' is called 'h1-eth0' and third interface of host 'h2' is called 'h2-eth2'. First port of switch 's1' is named 's1-eth1'. In switches, numbering begins with 1.

A. *Minimal*

Minimal is very simple topology that contains 1 OpenFlow switch and 2 hosts. It also creates links between switch and two hosts (Fig. 3).

# mn--topo minimal



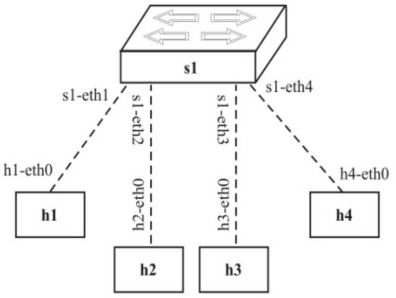
```
mininet> net
h1 h1-eth0:s1-eth1
h2 h2-eth0:s1-eth2
s1 lo:  s1-eth1:h1-eth0 s1-eth2:h2-eth0
c0
```

Fig. 3 Minimal Topology

B. *Single*

It is a simple topology with one openflow switch and k hosts. It also creates a link between switch and k hosts (Fig. 4).

# mn--topo single, 4

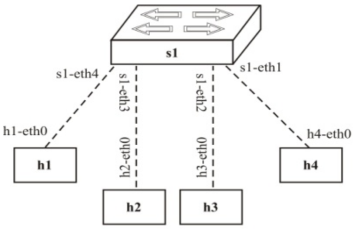


```
mininet> net
h1 h1-eth0:s1-eth1
h2 h2-eth0:s1-eth2
h3 h3-eth0:s1-eth3
h4 h4-eth0:s1-eth4
s1 lo:  s1-eth1:h1-eth0 s1-eth2:h2-eth0 s1-eth3:h3-eth0 s1-eth4:h4-eth0
c0
```

Fig. 4 Single Topology

C. *Reversed*

It is similar to single topology but connection order is reversed (Fig. 5).



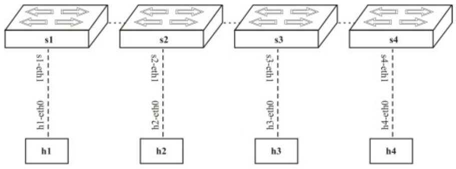
```
mininet> net
h1 h1-eth0:s1-eth4
h2 h2-eth0:s1-eth3
h3 h3-eth0:s1-eth2
h4 h4-eth0:s1-eth1
s1 lo:  s1-eth1:h4-eth0 s1-eth2:h3-eth0 s1-eth3:h2-eth0 s1-eth4:h1-eth0
c0
```

Fig. 5 Reversed Topology

# mn--topo reversed, 4

D. *Linear*

Linear topology contains k switches and k hosts. It also creates a link between each switch and each host and among the switches (Fig. 6).



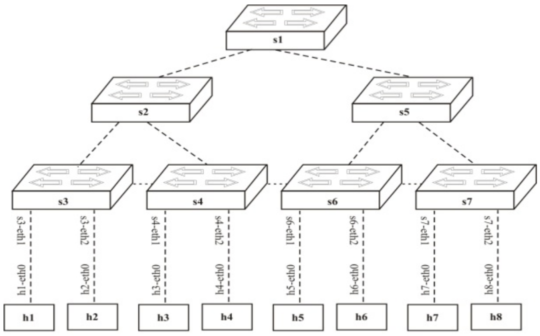
```
mininet> net
h1 h1-eth0:s1-eth1
h2 h2-eth0:s2-eth1
h3 h3-eth0:s3-eth1
h4 h4-eth0:s4-eth1
s1 lo:  s1-eth1:h1-eth0 s1-eth2:s2-eth2
s2 lo:  s2-eth1:h2-eth0 s2-eth2:s1-eth2 s2-eth3:s3-eth2
s3 lo:  s3-eth1:h3-eth0 s3-eth2:s2-eth3 s3-eth3:s4-eth2
s4 lo:  s4-eth1:h4-eth0 s4-eth2:s3-eth3
c0
```

Fig. 6 Linear Topology

# mn--topo linear, 4

E. *Tree*

Tree topology contains k levels and 2 hosts are attached to per switch (Fig. 7).



```
mininet> net
h1 h1-eth0:s3-eth1
h2 h2-eth0:s3-eth2
h3 h3-eth0:s4-eth1
h4 h4-eth0:s4-eth2
h5 h5-eth0:s6-eth1
h6 h6-eth0:s6-eth2
h7 h7-eth0:s7-eth1
h8 h8-eth0:s7-eth2
s1 lo:  s1-eth1:s2-eth3 s1-eth2:s5-eth3
s2 lo:  s2-eth1:s3-eth3 s2-eth2:s4-eth3 s2-eth3:s1-eth1
s3 lo:  s3-eth1:h1-eth0 s3-eth2:h2-eth0 s3-eth3:s2-eth1
s4 lo:  s4-eth1:h3-eth0 s4-eth2:h4-eth0 s4-eth3:s2-eth2
s5 lo:  s5-eth1:s6-eth3 s5-eth2:s7-eth3 s5-eth3:s1-eth2
s6 lo:  s6-eth1:h5-eth0 s6-eth2:h6-eth0 s6-eth3:s5-eth1
s7 lo:  s7-eth1:h7-eth0 s7-eth2:h8-eth0 s7-eth3:s5-eth2
c0
```

Fig. 7 Tree Topology

# mn--topo tree, 3

IV. Creating Custom Topologies

Using Mininet, you can easily create a custom topologies [11]. For example creating custom topology having 2 switches and 5 hosts (Fig. 8) needs just writing a few lines of Python [12] code. You can also easily create very complex flexible, robust. You can also configured that topology based on the parameters that are to be pass to it, and reuse that topology for multiple experiments.

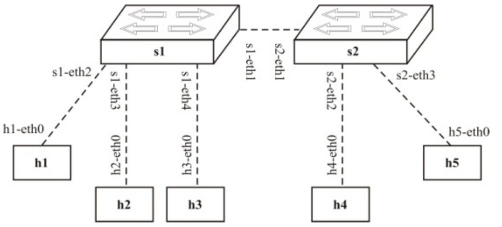
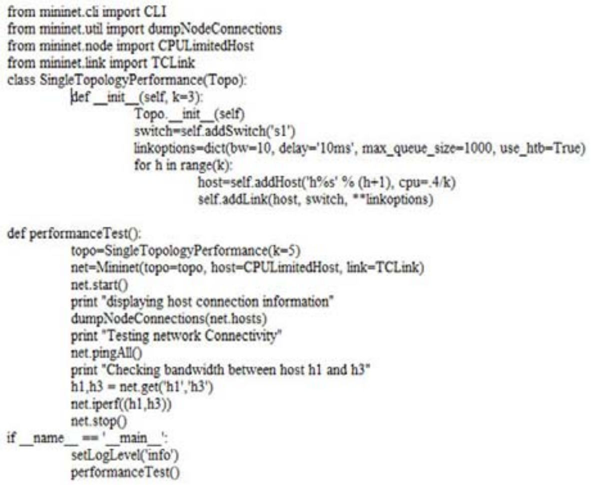


Fig. 8 Custom Topology

In the following Listing 1 contain Python code for creating custom topology having 2 switches and 5 hosts. This topology is run in Mininet by using following command.

# Python CustomTopologyPerformance.py



```
from mininet.cli import CLI
from mininet.util import dumpNodeConnections
from mininet.node import CPULimitedHost
from mininet.link import TCLink
class SingleTopologyPerformance(Topo):
        def __init__(self, k=3):
                Topo.__init__(self)
                switch=self.addSwitch('s1')
                linkoptions=dict(bw=10, delay='10ms', max_queue_size=1000, use_htb=True)
                for h in range(k):
                        host=self.addHost('h%s' % (h+1), cpu=.4/k)
                        self.addLink(host, switch, **linkoptions)

def performanceTest():
        topo=SingleTopologyPerformance(k=5)
        net=Mininet(topo=topo, host=CPULimitedHost, link=TCLink)
        net.start()
        print "displaying host connection information"
        dumpNodeConnections(net.hosts)
        print "Testing network Connectivity"
        net.pingAll()
        print "Checking bandwidth between host h1 and h3"
        h1,h3 = net.get('h1','h3')
        net.iperf((h1,h3))
        net.stop()
if __name__ == '__main__':
        setLogLevel('info')
        performanceTest()
```

Listing 1 Custom Topology Code

There are number of classes, functions, methods and variables in the Listing 1

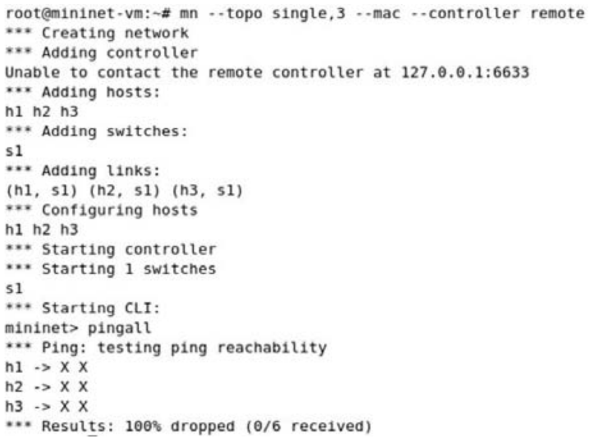
1. Topo: It is a base class for Mininet topologies.
2. Add Host (name, cpu = f): It is used for adding a host to the topology which contains two parameters. First parameter specifies the name of host and second specifies the fraction of overall system CPU resources that is to be allocated to the virtual host.
3. Add Switch(): It is used for adding a switch to the topology and returns the switch name, for example s1.
4. Add Link (node1, node2, **link options): It is used for adding a bidirectional link which contains three parameters. The first, second parameter specify the host and switch name respectively and third parameter specify the dictionary that contain number of options such as bandwidth, delay and loss characteristics, with a maximum queue size.
5. start(): It is used for starting your network.
6. stop(): It is used for stopping your network.
7. Mininet: It is used as a main class to create and manage a network.
8. net. hosts: It is used to show all the hosts in network.
9. ping All (): This is used to check connectivity between all nodes.
10. Set Log Level: There are number of Log Level such as info, debug, and output. Info is recommended as it provides useful information.
11. Dump Node Connections (): Dumps connections to/from a set of nodes.

There are basically two classes available such as CPU Limited Host and TC Link that can be used for performance limiting and isolation.

There are number of ways that these classes may be used, but simple way is to specify them as the default host and link classes to Mininet(), and then to apply the appropriate parameters in the topology.

V. Controlling Default Topology with DPCTL

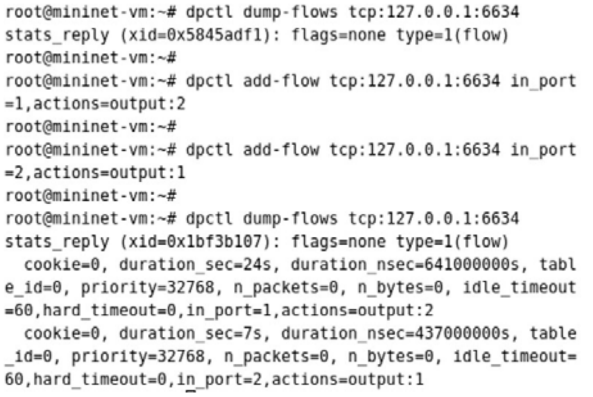
This topology (Fig. 9) creates 1 switch and 3 hosts. The option 'mac' set the mac address of each host according to node number and option 'remote' is to be used to connect switch to remote controller. As shown, hosts can not ping with each other because remote controller is not running, therefore flow table of switch does not contain any flow entry.



```
root@mininet-vm:~# mn --topo single,3 --mac --controller remote
*** Creating network
*** Adding controller
Unable to contact the remote controller at 127.0.0.1:6633
*** Adding hosts:
h1 h2 h3
*** Adding switches:
s1
*** Adding links:
(h1, s1) (h2, s1) (h3, s1)
*** Configuring hosts
h1 h2 h3
*** Starting controller
*** Starting 1 switches
s1
*** Starting CLI:
mininet> pingall
*** Ping: testing ping reachability
h1 -> X X
h2 -> X X
h3 -> X X
*** Results: 100% dropped (0/6 received)
```

Fig. 9 Single Switch, 3 Hosts Topology

There are two methods to add flow entries into flow table of switch, remote controller and 'dpctl' [13] utility that works on port 6634 (Fig. 10). dpctl is a data path controller that comes with OpenFlow reference distribution and is used to manage the flow table of switch by using 'dpctl' commands. After adding flow entries by using ‘dpctl’ command, the host h1 and h2 can ping with each other (Fig. 11).

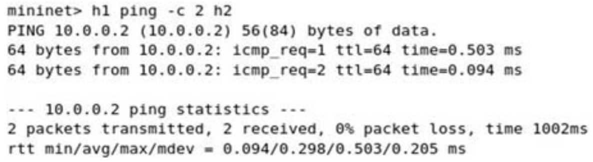


```
root@mininet-vm:~# dpctl dump-flows tcp:127.0.0.1:6634
stats_reply (xid=0x5845adf1): flags=none type=1(flow)
root@mininet-vm:~#
root@mininet-vm:~# dpctl add-flow tcp:127.0.0.1:6634 in_port
=1,actions=output:2
root@mininet-vm:~#
root@mininet-vm:~# dpctl add-flow tcp:127.0.0.1:6634 in_port
=2,actions=output:1
root@mininet-vm:~#
root@mininet-vm:~# dpctl dump-flows tcp:127.0.0.1:6634
stats_reply (xid=0x1bf3b107): flags=none type=1(flow)
  cookie=0, duration_sec=24s, duration_nsec=641000000s, tabl
e_id=0, priority=32768, n_packets=0, n_bytes=0, idle_timeout
=60,hard_timeout=0,in_port=1,actions=output:2
  cookie=0, duration_sec=7s, duration_nsec=437000000s, table
_id=0, priority=32768, n_packets=0, n_bytes=0, idle_timeout=
60,hard_timeout=0,in_port=2,actions=output:1
```

Fig. 10 Flow Management Using DPCTL

Default idle timeout of each flow entry is 60s. Therefore after 60 seconds, flow entries will expire and needs to be added again and again. So installs a flow entry with longer timeout using following command:

dpctl add-flow tcp:127.0.0.1:6634 in_port = 1, idle_timeout = 180, actions = output:2



```
mininet> h1 ping -c 2 h2
PING 10.0.0.2 (10.0.0.2) 56(84) bytes of data.
64 bytes from 10.0.0.2: icmp_req=1 ttl=64 time=0.503 ms
64 bytes from 10.0.0.2: icmp_req=2 ttl=64 time=0.094 ms

--- 10.0.0.2 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1002ms
rtt min/avg/max/mdev = 0.094/0.298/0.503/0.205 ms
```

Fig. 11 Connectivity Test

VI. Conclusion

Mininet is a platform for rapid network prototyping. It can run unmodified network application code on small networks as well as very large networks. It is an alternative to run SDN experiments on emulated networks. Real systems are very painful to reconfigure. Virtual machines allow easier topology changes but suffer from scalability issues. Simulators are a good alternative but same source code cannot be deployed on real hardware. There are performance issues on Mininet. The challenges before Mininet are to model networks of very large scale with practical performance.

Acknowledgment

The authors would like to thank Mr. Vipin Gupta from U-Net Solutions, Moga, India for his valuable help.

References

[1] Handigol, Nikhil, Brandon Heller, Vimal kumar, Jeya kumar, Bob Lantz, and Nick McKeown. "Reproducible network experiments using container-based emulation." In *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, pp. 253–264. ACM, 2012.

[2] Lantz, Bob, Brandon Heller, and Nick McKeown. "A network in a laptop: rapid prototyping for software-defined networks." In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, pp. 19. ACM, 2010.

[3] POX at https://openflow.stanford.edu/display/ONL/POX+Wiki# POXWiki-forwarding.l2_learnining.

[4] Feamster, Nick, Jennifer Rexford, and Ellen Zegura. "The road to SDN: an intellectual history of programmable networks." *ACM SIGCOMM Computer Communication Review* 44, No. 2 (2014): 87–98.

[5] Nunes, B.; Mendonca, M.; Nguyen, X.; Obraczka, K.; Turletti, T., "A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks," *Communications Surveys & Tutorials, IEEE*, Vol. 1, No.99, pp. 18.

[6] Shenker, Scott, M. Casado, T. Koponen, and N. McKeown. "The future of networking, and the past of protocols." *Open Networking Summit* (2011).

[7] McKeown, Nick, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. "OpenFlow: enabling innovation in campus networks." *ACM SIGCOMM Computer Communication Review* 38, No. 2 (2008): 69–74.

[8] Fernandez, Marcial P. "Comparing openflow controller paradigms scalability: Reactive and proactive." In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, pp. 1009–1016. IEEE, 2013.

[9] Fernandez, Marcial. "Evaluating OpenFlow Controller Paradigms." In *ICN 2013, The Twelfth International Conference on Networks*, pp. 151–157. 2013.

[10] Mininet Topologies at http://www.routereflector.com/2013/11/mini net-as-an-sdn-test-platform

[11] Mininet at https://github.com/mininet/mininet/wiki/Introduction-to-mininet.

[12] Python at https://www.python.org/.

[13] dpctl at http://archive.openflow.org/wk/index.php/HOTITutorial2010.

Time Synchronization Strategies in Wireless Sensor Network: A Review

Ekta[1] and Jyoteesh Malhotra[2]
[1,2]Department of Computer Science,
Guru Nanak Dev University, Regional Campus, Jalandhar, India
E-mail: [1]er_ekta@yahoo.com, [2]jyoteesh@gmail.com

***Abstract*—Lately, Wireless Sensor Networks are getting lot of attention by research community. It consists of small devices distributed over geographical areas and each device has sensing, computing and communicating components. Time synchronization is vital to schedule communication and distributed measurement tasks. It targets at equalizing the local times for all nodes in the network. It has been observed from the literature survey that the task of time synchronization mainly focus on two things, firstly drift from master oriented towards master less synchronization and secondly energy efficiency techniques. Various techniques purposed in the literature have been thoroughly covered in this survey paper and related issues and challenges have been highlighted.**

Keywords: Wireless Sensor Network, Time Synchronization, Energy Efficiency

I. Introduction

Wireless sensor network is a network which is consisting of distributed autonomous sensors to monitor physical or environmental conditions, such as temperature, sound, pressure etc. and to cooperatively pass their data through the network to a main position [10]. It consists of nodes which vary from a few to several hundreds or even thousands, where each node is linked to one (or may be several) sensors. The main components of sensor node are a microcontroller, transceiver, external memory, power source and one or more sensors [16]. The controller performs tasks, processes data and controls the functionality of other components in the sensor node. The functionality of both transmitter and receiver is combined into a single device known as a transceiver. Memory requirements are very much application dependent. An important aspect in the development of a wireless sensor node is ensuring that there is always adequate energy available to power the system. The sensor node consumes power for sensing, communicating and data processing.

Time synchronization means to scale all nodes of network towards common clock time so that communication is effective. Since all hardware clocks are imperfect, those at different nodes may drift away from each other in time. For this reason, the observed time or durations of time intervals may differ for each node in the network. However, for many applications or networking protocols, it is required that a common view of time exists and is available to all or some of the nodes in the network at any particular instant. Many strategies are available having their respective pros and cons. The diversity between strategies is basically on the factors like using master node or not, energy efficiency, scheme used for waking and sleeping time of nodes etc. The traditional basic protocol is Two Way Message Exchange Synchronization [6]. It is base of many protocols which are described in this paper.

Moreover, rest of the paper is organized as: Section II describes the basis of time synchronization. Section III presents various time synchronization strategies. Based on the literature survey, the important issues of time synchronization have been discussed in Section IV, before the paper is finally concluded in Section V.

II. Backbone of Time Synchronization

Time synchronization in wireless sensor network has basic requirements as computer clock, knowledge of sources of synchronization errors and needs of synchronization algorithms [6]. These basic requirements are as discussed below:

A. *Computer Clock*

Computer clock circuits consist of an oscillator and a counter. Depending on the oscillator’s angular frequency, the counter increases its value to represent the local clock C(t) of a network node. In ideal situations, angular frequency is constant. The angular frequency changes and computer clocks drifts due to physical variations, like temperature, vibration and pressure. Approximation of local clock can be made as [11]:

$$C_i(t) = a_i t + b_i \qquad (1)$$

Where a_i is the clock drift, and b_i is the offset of node i's clock. Drift represents the rate of the clock, and offset means the difference in value from real time t. Using (1), local clocks of two nodes in a network can be compared, say node 1 and node 2 as in [12]:

$$C_1(t) = a_{12.}\, C_2(t) + b_{12} \qquad (2)$$

We call a_{12} the relative drift and b_{12} the relative offset between the clocks of node 1 and node 2. If two clocks are perfectly synchronized, then their relative drift is 1, meaning the clocks have the same rate. Their relative offset is zero, meaning they have the same value at that instant. Few strategies are designed to adjust offsets of nodes.

B. *Sources of Synchronization Errors*

The following elements contribute to the synchronization errors [6]:

1. Send time which is the total time of building the message and transfer it to the network interface to

be sent. This time highly depends on the operating systems in use.

2. Access time which is the time needed to access the channel. Every network employs a medium access control (MAC) scheme.
3. Propagation time which is the time required to propagate the message through the air from network interface of the sender to the network interface of the receiver.
4. Receive time which is the time spent in receiving the message by network interface and transferring it to the application layer of the host.

C. *Needs of Synchronization Algorithms*

While designing time synchronization algorithm, wireless sensor network limitations enforce certain requirements that need to be met [8]. These are accuracy, robustness, scalability, longevity, energy efficiency, cost, scope, delay which has been briefly discussed below:

1. *Precision:* It means how accurately the algorithm is performing its task which varies from application to application.
2. *Robustness:* In case of harsh environment, synchronization scheme should be functional.
3. *Scalability:* Synchronization scheme should tolerate the change in the number of nodes in the network.
4. *Longevity:* It means synchronization scheme must be able to work as long as operation of the network.
5. *Energy Efficiency:* Energy is very critical part of node in wireless sensor network so synchronization scheme must conform to energy efficiency.
6. *Cost:* Sensor nodes are very small and inexpensive devices so attaching high cost hardware for synchronization process do not make any sense.
7. *Scope:* The time synchronization scheme aims at common time scale for all nodes in the network. Network can be global or local but its difficult to implement such algorithms in case of global network.
8. *Delay:* Time synchronization must be done in small time as possible for some critical applications such as gas leak detection.

III. Synchronization Strategies for Wireless Sensor Network

A. *Reference Broadcast Synchronization for Wireless Sensor Network*

To get rid of overhead of sender in sender-receiver synchronization, receiver-receiver synchronization was proposed named as Reference Broadcast Synchronization [16] which concentrate on synchronizing a set of receivers with one other rather than synchronizing a sender with a receiver. In this scheme, nodes send reference beacons to their neighbors. A reference beacon does not include a timestamp. Instead, its time of arrival is used by receiving nodes as a reference point for comparing clocks. In this, reference sender remains unsynchronized because in doing so more energy is wasted.

B. *Lightweight Tree-Based Synchronization*

Lightweight Tree-Based Synchronization [14] is based on construction of tree which is constructed each time whenever synchronization is to be performed. Once the tree has been constructed, the reference node initiates pair-wise synchronization with each of its children. Then each child repeats this step for its respective children and this repetitive process is performed until all nodes of the tree have been synchronized. It aims at attaining minimum complexity for providing a given precision.

C. *Time-Sync Protocol for Wireless Sensor Network*

Lightweight tree based synchronization aims at a particular precision rather than very high precision which may not be suitable for some applications. For achieving higher precision Time-sync [13] was proposed. It is a sender-receiver synchronization approach. It has two steps to be implemented. First step is Level Discovery and second is Synchronization. Level Discovery focuses on creating the hierarchical topology in the network, where each node is assigned a level. Where only one node is assigned level 0 which is known as root node. In the second step, node at level i synchronizes with node at level i-1. At the end of hierarchy, all nodes in the network are synchronized to root node. It has two times more precision than Reference Broadcast Synchronization. Its shortcoming is less accuracy as it does not estimate the clock drift of nodes.

D. *Flooding Time Synchronization in Wireless Sensor Network*

For providing a robust synchronization against link or node failure, Flooding Time Synchronization [9] abstracts the ideas of both RBS (Reference Broadcast Synchronization) and TPSN (Time-sync Protocol). It enhances the key ideas of TPSN and RBS and combines them to periodically flooding the synchronization messages to achieve network wide synchronization which is robust against link or node failure. It is a sender-receiver synchronization as in case of TPSN and also adjusts the clock drift as in case of RBS. In FTSP, all nodes in the network synchronize to a dynamically (re) elected root node by the use of controlled flooding. It uses low communication width. Its lackings are high energy consumption and poor synchronization of distant nodes.

E. *Recursive Time Synchronization Protocol*

To overcome the drawbacks of Flooding Time Synchronization, Recursive Time Synchronization [3] was proposed which targets at better performance than the earlier protocols. This protocol attains time synchronization by exchanging messages among nodes. Three types of messages are used which are enquiry/election of reference node, request for time synchronization and reply for time synchronization. When wireless sensor network boots up, all sensor nodes need to identify the reference node. For this purpose, these nodes sends enquiry message to their respective neighbors and wait for some time until reply is received. The node which has the smallest ID is chosen as reference node. This protocol takes responsibility of two things which are choice of reference node and compensation of offset and drift. This protocol uses seven times less energy than Flooding Time synchronization protocol in long run. By compensating propagation delay, accuracy is improved. Energy efficiency is achieved by using techniques like infrequent broadcasts by reference node, reducing number of synchronization requests.

F. *Mini-Sync and Tiny-Sync*

In order to attain synchronization protocol which is more constrained to communication bandwidth, Mini-Sync and Tiny-Sync[7] was proposed which aims at minimal computational and storage complexity. In this algorithm, concept of data points is used which is the tuple of timestamps. Tiny-sync is based on the concept that all data points are not useful so it applies constraints on the selection of data points. It uses low storage space as only four data points and eight time stamps are to be stored. Mini-Sync improves the accuracy of Tiny-Sync at a low computational cost as it discards only that constraint if newer constraint can eliminate the existing constraint. These protocols are suitable for sensor networks that are highly constrained in bandwidth and computational power providing tight, deterministic synchronization. No scalability and robustness is defined for these protocols. It is also tolerant of message losses.

G. *Master–Less Time Synchronization for Wireless Sensor Networks*

Above discussed protocols mostly uses tree based organization which includes the overhead of selecting a master node so as to synchronizing its respective children. To reduce such overhead, Master-Less time synchronization strategy [5] aims at scaling all network nodes to common clock without using any master node. All nodes periodically broadcast their own timestamps in a pseudo-periodic manner and adjust their own clock as soon as they receive the time from some other device. While working in all conditions, convergence speed and accuracy of course strongly depend on the specific scenario considered. Its main advantages are reliability and flexibility, since no special master election procedures or multi-hop synchronization protocols are required.

H. *Efficient Time-Synchronization in Ring-Topology Wireless Sensor Networks*

Concentrating on energy efficiency of synchronization protocols, efficient time-synchronization [4] in ring topology was designed for low duty cycle networks sensors which sleep most of the time and wake up for short periods of time to perform sensing, computing, and communication tasks. This is based on use of two software clocks for synchronization purpose. The first software clock is called Wake-up Clock (WUC). This clock controls the wake-up of the nodes for transmission and reception, according to the pre-determined schedule. The other clock is called Sync Clock (SC), and it contains the actual synchronization information. Most of the time the two clocks show the same time but there are time intervals when they diverge. Initially both clocks show the same time. The node is woken up by the WUC to receive a message. The message contains new time stamp and the node adjusts its SC accordingly. The node then transmits a message, using the content of SC for time stamping. After message transmission, but before the next message reception, the WUC is adjusted to SC, thus from now on both clocks show the same time, until the next message reception.

I. *Enhanced Time Synchronization Algorithm*

As network topology is not fixed, it can be ring, tree, star or any other so time synchronization with energy efficiency is required for varying network topologies and for the same reason, Enhanced Time Synchronization algorithm [2] was proposed which focuses at reducing the power consumption in the distributed network. It has mainly two parts: first is synchronization time and second is energy dissipation. In this synchronization in the network takes place by autonomous and local decisions of sensor node. There is no centralized control. In synchronization phase, node with smallest ID is selected as transmitter after node deployment. This selected node sends syn_msg to next node ID which in turn acts as receiver node and sends syn_ack back to transmitter and hence information exchange takes place. Now receiver node becomes the transmitter node and next node ID is selected a s receiver node. This process is repeated until all nodes get synchronized. After this, energy dissipation is calculated. Synchronization time slightly increases with increase in number of nodes.

J. An Energy-Efficient Beacon Strategy for Time Synchronization in Wireless Sensor Network

As Efficient time-synchronization in ring-topology [3] uses the concept of low-duty cycle, means including the concept of waking/sleeping modes of nodes, for attaining energy efficiency which has a significant drawback of missing synchronizing beacon due to wrong active/sleep mode of node due to clock drift. So for overcoming this drawback, An Energy-Efficient Beacon Strategy [1] was developed which is based on guard beacon. Using guard beacon, a node will send optimally scheduled beacons to minimize the idle listening time for the receiver caused by synchronization errors. It aims at reducing overall power consumption of both sender and receiver node rather than just reducing the number of synchronization messages. By sending the optimal number of beacons at the optimal times, Guard Beacon can reduce the idle listening time for coming beacon, therefore minimize the total synchronization power consumption of sending and receiving beacons. This protocol has more energy efficiency than Recursive time synchronization protocol [3] when the beacon interval is small.

All above surveyed time synchronization protocols are briefly highlighted in Table 1. In this, overview of each protocol is mentioned which includes accuracy, network topology, energy efficiency etc.

TABLE 1 VARIOUS TIME SYNCHRONIZATION STRATEGIES AT A GLANCE

Sr. No.	Strategy Name	Highlighted Points
1.	Reference Broadcast Synchronization	i. Receiver-Receiver Synchronization ii. Removes Non-Deterministic Synchronization Errors Caused by Sender iii. More Energy Wasted if Reference Node to be Synchronized
2.	Light-Weight Tree based Synchronization	i. Aims at Reasonable Accuracy ii. Based on Tree Structure iii. Accuracy Decreases as Tree Depth Increases iv. Sender-Receiver Synchronization
3.	Time-Sync Protocol	i. Works in Two Phases: Level Discover Phase and Synchronization Phase ii. Based on Tree Structure iii. Sender-Receiver Synchronization iv. Error Depends on Complexity of the Tree Structure
4.	Flooding Time Synchronization	i. Uses Ideas of Both Time-Sync Protocol and Reference Broadcast Synchronization ii. Utilizes Low Communication Bandwidth iii. Sender-Receiver Synchronization iv. Frequent Broadcast of Synchronization Messages
5.	Recursive Time Synchronization	i. Sender-Receiver Synchronization ii. Improves Energy Efficiency by Using Infrequent Broadcast of Synchronization Messages iii. Improves Accuracy by Compensating Propagation Delay iv. Better Performance than Flooding Time Synchronization
6.	Mini-Sync and Tiny-Sync	i. Provide Tight, Deterministic Synchronization with Low Storage and Computational Complexity ii. Tolerant Against Message Failure iii. No Scalability and Robustness is Defined
7.	Master–Less Time Synchronization	i. No Need of Election of Master Node ii. The Algorithm Converges within Reasonable Time Regardless of Time-Stamping Jitter, Incomplete Node Visibility
8.	Efficient Time-Synchronization in Ring-Topology	i. Proposed for Low Duty Cycle Ring Topology ii. Very Low Synchronization Error for Neighbors iii. Saves Power as Most of the Time Nodes are in Sleeping Mode
9.	Enhanced Time Synchronization Algorithm	i. Takes Less Time for Synchronization ii. Reduces Power Consumption iii. Synchronization Time Slightly Increases with Increase in Number of Nodes
10.	An Energy-Efficient Beacon Strategy	i. Sends Multiple Beacons for Synchronization ii. Reduce Power Consumption of Both Sender and Receiver

…Table 1 (Various Time Synchronization Strategies at a Glance)

IV. OPEN ISSUES

Based on the extensive survey done in the previous section, a lot of work has been reported in the literature. However, there is still a need of comprehensive time synchronization techniques that covers all requirements holistically. However, main issues in this have been highlighted here that includes send time, propagation time, receive time, access time which results in delay. If delay occurs in a network than clock times are not properly synchronized. Many protocols, like Reference broadcast synchronization, Flooding time synchronization have sorted out these issues so that the delay can be negligible. Another issue is energy efficiency which is very vital factor in wireless sensor network. Some protocols, like Efficient time-synchronization in ring-topology, Energy-efficient beacon strategy, exists which conform to energy efficiency to some extent but more energy efficient and high precision is still a big issue.

V. CONCLUSION

Due to challenges associated with Wireless sensor network such as energy efficiency, self-configuration, size and sensor mobility, time synchronization in wireless sensor network is more challenging as compared to other wireless networks. Time synchronization is needed to identify casual relationships between events in physical world. In this paper, time synchronization strategies have been reviewed and major issues have been identified that mainly includes delay and energy efficiency. From the survey, it has been observed that researchers have considered the solution of time synchronization from two aspects that are viz. centralized or distributed control and in the energy constraint. From the literature, it can be concluded that design innovations are needed to develop time synchronization strategies for energy efficiency and improved precision for hostile and challenging wireless sensor network.

REFERENCES

[1] Yongrui Chen, Fei Qin, and Weidong Yi," Guard Beacon: An Energy-Efficient Beacon Strategy for Time Synchronization in Wireless Sensor Network", IEEE Communications Letters, Vol. 18, No. 6, June 2014.

[2] Kaushal R., Buttar A.S.," An Alternative Approach for Energy Efficient Time Synchroniaztion in Wireless Sensor Networks", International Journal of Advanced Research in Computer Engineering & Technology Vol. 3, Issue 3, March 2014.

[3] M. Akhlaq and T.R. Sheltami, "RTSP: An accurate and energy-efficient protocol for clock synchronization in WSNs," IEEE Trans. Instrum. Meas., Vol. 62, March. 2013.

[4] Simon, "Efficient time-synchronization in ring-topology wireless sensor networks", IEEE International Instrumentation and Measurement Technology Conference, 2012, pp. 958–962.

[5] Daniele Fontanelli, David Macii," Master–less Time Synchronization for Wireless Sensor Networks with Generic Topology", Instrumentation and Measurement Technology Conference, 2012, pp. 2785–279.

[6] Sami M. Lasassmeh and James M. Conrad, "Time Synchronization In Wireless Sensor Network: A survey" IEEE SoutheastCon 2010, pp. 242–245.

[7] S. Yoon, C. Veerarittiphan, and M. Sichitiu, "Tiny-sync: Tight time synchronization for wireless sensor networks," ACM Trans. on Sensor Networks (TOSN), Vol. 3, No. 2, pp. 1–33, Jun. 2007.

[8] F. Sivrikaya and B. Yener, "Time Synchronization in Sensor Networks: A Survey," Network, IEEE Vol. 18, Issue 4, July-Aug. 2004, pp. 45–50.

[9] M. Maro`ti, B. Kusy, G. Simon, and A. Ldeczi, "The flooding time synchronization protocol," in Proc. of the 2nd ACM Conference Embedded Networked Sensor Systems, Baltimore, Maryland, USA, Nov. 2004, pp. 39–49.

[10] K. Römer and F. Mattern. "The Design Space of Wireless Sensor Networks," Wireless Communications, IEEE Vol. 11, Issue 6, Dec. 2004, pp. 54–61.

[11] S. Ping, "Delay Measurement Time Synchronization For Wireless Sensor Networks," Technical Report, Intel Research Berkeley Lab, 2003.

[12] M.L. Sichitiu and C. Veerarittiphan, "Simple, Accurate Time Synchronization for Wireless Sensor Networks, "Wireless Communications and Networking, IEEE Vol. 2, March 2003, pp. 1266–1273.

[13] S. Ganeriwal, R. Kumar, and M.B. Srivastava, "Timing–sync protocol for sensor networks," in Proc. of the 1st ACM Conference on Embedded Networked Sensor Systems, Los Angeles, California, USA, pp. 138–149, Nov. 2003.

[14] Van. Greunen J., Rabaey J, "Lightweight time synchronization for sensor networks", Proc. Of the International Workshop on Wireless Sensor Networks and applications, 2003.

[15] J. Elson, L. Girod, and D. Estrin, "Fine-Grained Time Synchronization using Reference Broadcasts," Proceedings of the Fifth Symposium on Operating Systems Design and Implementation, Boston, MA, December 2002.

[16] J. Hill and D. Culler, "A Wireless Embedded Sensor Architecture for System-Level Optimization," Technical Report, U.C. Berkeley, 2001.

Applications of Data Mining – A Survey

Lovedeep[1] and Sabia[2]
[1,2]Department of Computer Sc. & Engineering,
Guru Nanak Dev University Regional Campus, Jalandhar, India
E-mail: [1]Lovedeep1991@gmail.com, [2]sabiajal@gmail.com

***Abstract*—Data mining is an analytic process designed to explore large amount of data and to predict new subset of data. There are different MNC's and organizations that operate at different places in different countries and generate massive data at each operation place. The decision making process requires to gather all data and with help of different data mining techniques the useful information is abstracted from large data items. In this paper, the main focus is how data mining is useful in different organizations such as retail stores, banks, hospitals and insurance companies where we have to deal with massive data and list the applications of data mining in different fields.**

Keywords: Data Mining Model, Decision Making Using Data Mining, Stages, Data Mining, KDD

I. INTRODUCTION

Now a day people around the world are using vast data in different fields such as health care, business, industries etc. and this data is not in same format everywhere, we are having different forms of data such as documents, graphics, audio formats etc. Data mining is basically retrieving useful information from these huge databases. This technique is also called KDD (knowledge discovery process). We are having large amount of data but there is lack of useful information that is "we are data rich but information poor". It is very hard process to turn large data volumes into useful and worth decision making knowledge. To solve this problem of retrieving useful information from large set of data, we are moving toward the field of data mining. To take maximum advantage of raw data only its retrieval is not enough. It require different tools and methods for automated submission of data, extraction of useful items and information from it, discovery and prediction of new patterns and interpretation of data for decision making process. This all can be done efficiently by "data mining".

Data mining is to extract hidden predictive information from large databases and it helps organizations to focus them on most important information of their warehouses [1, 2]. Data mining tools are useful in predicting the future trends and behaviours which are helpful for organizations to take proactive knowledge driven decisions. It is also saves our time as data is reduced after information is exacted from it. The data mining process consist of three major stages are:

1. Initial exploration: Exploration means to search or investigate. In this stage, the data preparation is done which involve cleaning data, data transformations, selecting subsets of records. In case of having large number of variables or fields in sub sets, we perform some preliminary feature operations to keep variables in a manageable range.
2. Model building and validation: In next stage the relevant variables and complexity is determined. Various models are considered bases on the predictive performance. It is very long process and various techniques are used to compare different models and chose best one called competitive evaluation of models.
3. Deployment: The model which is selected in the last stage is as best model is used and new data values are applied to it for future estimations and predictions.

This paper is organised as follows. In section II literature survey. In Section III the various applications has been discussed. Final section gives a conclusion of the paper.

II. LITRATURE SURVEY

1. Dominik Fisch [7] in 2014 author introduce new way of fusing classifiers at the level of parameters of classification rules the two novel technique to fuse probabilistic generative classifiers(CMM) into one. This technique based on multinomial distributions for categorical input and multivariate normal distributions. The main advantage of this fusion is to hyper distributions throughout the fusion process used components used in online training.
2. Jianlin Xu, [8] in 2013 to improve the security status of mobile app author propose a methodology for mobile apps based on data mining and cloud computing to filter out malware apps from mobile app markets and also present prototype system is mobsafe. Mobsafe combined the static and dynamic analysis methods are SAAF and ASEF to estimate the total time needed to evaluate all the apps stored in one mobile app market.
3. Neelamadhab Padhy, [16] in 2012The author talks about the two major applications of data mining i.e. generic applications and domain specific applications. It is observed that no generic application is fully generic. There are limitations of generic applications of data mining. Domain and data, context parameters and aim of data mining try to influence the data mining decisions. Domain specific applications produce more accurate results which are over 90% and these are more specific for data mining. It is difficult to design such mining system which works for any domain dynamically.

4. Shouyi Wang [4] in 2011 author describes that numerical errors are detectable in advance; data mining techniques can be used for early detection of these numerical typing errors. Author used multichannel electroencephalogram (EEG) recordings for quantitative analysis of detecting errors along with two basic data mining techniques i.e. linear discriminate analysis (LDA) and support vector machine (SVM). Around 80% errors can be detected even without processing data saving lot of time. Using data mining techniques it is possible to proactively predict the keystrokes with errors based on EEG recordings. The only drawback is that the study is based on limited data pool, it maybe not produce result in generalized form.
5. Mahdi Esmaeili [5] in 2010 author presented there is huge number of variables and objectives involved in aerospace engineering optimization which cannot be ignored, so only multi-objective optimization can deal with it effectively and also proposed method to use data mining techniques for aerospace engineering optimization process. The advantage of using these techniques is that less variables decrease the effective cost of optimization. Here simple variable reduction tool and some data mining techniques are applied for getting desired results. Three algorithms are used which speed up the optimization process. Even in worse case 55% variables are reduced and BFTree and J48 algorithms require fewer variables while LAD Tree algorithm utilizes at least 7 variables to classify data set.
6. Sérgio Ramos, Zita Vale [6] in 2008 author present an electricity medium voltage (MV) consumers characteristics and classifications and also compare the three different clustering algorithms for taking the number of clusters and it also presents the new tariff structures to apply for each customer class.
7. Bartley D. Richardson [9] in 2008 author describes the techniques used by fellow's reactions and feedback from the students by integrates data mining and software engineering. STEP project is based on engineering principles and relevant to student's lives. To increase the confidence and learning abilities of students using STEM subject's mega miming mart is designed to understand correlations in a familiar setting and periodic table.
8. Elovici, Y., Kandel [12] author proposed the knowledge based methodology to view terror-related context at a series of evasive web sites. This methodology use data mining algorithm to the textual context of terror-related web sites. User can view all activities done by terrorist. Author also describes the intrusion detection system (IDS), vector space model and clustering techniques.

III. Applications

A. *Medical Science and Health Care Field*

There is constant work going on in field of healthcare and medical sector to improve the facilities for patients. At same time it is hard to deal with data captured around healthcare processes in form of electronic health records, health insurance claims, disease registries and clinical trials. Data getting huge day by day and task here is to not only store data but use it effectively so as to improve medical facilities. Data mining can be used for this purpose as medical data is very complex and difficult to analyze. The data mining algorithms are helpful in reducing patient's risks and diagnosis costs. Using the prediction algorithms the observed prediction accuracy was 100% for 91.3% cases [3]. Also the success in health care depends upon the availability of clean data. Data should be capture properly, stored in proper manner for better mining of data.

B. *Market Analysis and Retail Stores*

With market segmentation, it is easy to find out behaviours that are same among the customers. You can look for patterns among customers that seem to purchase the same products at the same time. Another application of data mining is called customer churn. Customer churn helps in estimating that which customers are likely to stop purchasing products or services from you and go to one of your competitors. Also any company can use data mining to find out which purchases are the most likely to be fraudulent. Using data mining a retail store may be able to find out the products which are stolen maximum number of time. By finding out which products are stolen the most, steps can be taken to protect those products and detect those who are stealing them. Sequential pattern mining is effective mining method used in field of marketing.

C. *Predictions in Engineering Field*

In recent years, there is increase in the use of data mining techniques on such artefacts are with the goal of analyze and improving software processes for a given organization or project. As data mining provides techniques like intentional mining for predictions of data, it can be used in field of engineering. We can use data mining for predictions of cost estimation in engineering field.

D. *Network Security*

Security within internet is very serious issue. The e-business culture is developing more and more in recent years as effect there is more threat posed by internet crime. Data mining techniques such as association rule and clustering analysis are quite useful in preventing the network from these threats. Data mining utilize the light weight statistical summaries which are gathered at distributed points within a network for detecting the security threats with help of signature detection filtering mechanism [10].

E. The Intelligence Agencies

Data mining is quite useful in field of intelligent agencies. These agencies have to deal with terrorists and other threats to countries. With help of data mining, agencies collect data regarding the activities of terrorists to investigate their future attacks. The Clustering technique is helpful (Association rule mining) for the different objects (like persons, organizations, vehicles etc.) in crime records. Data mining detects as well as analyzes the crime data. Agencies are developing new algorithms which make it easy to work with large data set [12].

F. Data Mining in Field of Sports

The world of sports produces a huge statistical data about each team and player. This data can be very useful for the assist coaches and managers if it can be managed properly by experts. Now a day sports data mining is very popular trend for prediction of result and player assessment. Even it can be utilized for identification of new talent, predictions about injuries and game strategies. By using data mining tools the experts can help players to improve the performance. Data mining techniques such as ANN, decision trees, Bayesian method, SVM, logistic regression and fuzzy methods have been employed to predict game results. By using hybrid algorithms the accuracy of prediction can be increased [11].

G. Web Education Systems

The courseware can be improved by using data mining techniques in web education. Data can be picked from student sessions on web and can be utilized by teachers or authors to improve the content and effectiveness of course. Data mining help increasing awareness among the learners [13].

H. Broadcasting Data in Mobile Computer Scenario

By sing wireless communication devices, mobile can be accessed from any place. As we know downlink speed is more as compared to uplink speed, we can use data mining techniques in broadcasting of data. There are two methods used for this purpose. We try to arrange the subsequently requested data items are places closed to each other. It is also seen that using data mining in broadcasting data, the access latency can be decreased very efficiently [14].

I. Financial and Investment Field

Economic growth is one of the major issues for developing countries like India. Information technology plays great role in economic growth and globalization of economy. Due to this vast financial data being generated day by day and accumulated at unprecedented place. Data mining provides automated approach for utilization of vast data and to give back information to financial companies for better investment strategies. Hidden patterns are discovered by data mining techniques like Sequential pattern and time-series mining, clustering analysis and association rules [15].

IV. Summary

TABLE 1

Sr. No	Data Mining Methods and Techniques	Applications Field	Type of Data Set
1.	Clustering Technique	Bio-Informatics, Medical Sciences, Marketing and Business, World Wide Web	Statistical Data, Discrete and Comparative Numerical Data.
2.	Decision Tree	Operational Research, Field of Sports, Sales and Purchase	Graphical Data, Trees and Graphs
3.	Sequential Pattern Mining	Retail Marketing, Financial and Investment Field	Continuous Data, Numeric or Alphanumeric
4.	Intentional Mining	Web Search, Business, Engineering	Interactive Data from Device Such as Computers, Images etc
5.	Association Rule	Intelligence Agencies, Security of Networks	Large Variable Data Set, Numerical, Alphanumeric, Visual and Audio Data
6.	Factor Analysis	Field of Psychology, Intelligencer Verbal Intelligence	Statistical Data, Large and Numerical values Visual Data Sets

V. Conclusion

The paper provided various applications of data mining and the mining methods that are used in different fields. As we know that data varies such as text, graphical and audio data in different fields so it is important to deploy a method which reduce variable in reasonable way by using data mining tools and methods such as clustering and sequential pattern mining. Depending upon the data type available in a particular field, method is selected to make data mining process convenient and swift. Evolving data mining applications have shown great potentials in financial, medical, network security and engineering fields and will continue to prosper in new fields such as sports and web education. By using proper mining techniques more applications can be explored in future with higher efficiency.

References

[1] Larose, D.T., “Discovering Knowledge in Data: An Introduction to Data Mining” 0-471-66657-2, John Wiley & Sons, Inc, 2005.

[2] Dunham, M.H., Sridhar S., “Data Mining: Introductory and Advanced Topics”, Pearson Education 81-7758-785-4, 1st Edition, 2006.

[3] Neelamadhab Padhy, Dr. Pragnyaban Mishra, Rasmita Panigrahi " Survey of Data Mining Applications And Feature Scope" IJCSEIT, (Vol. 2) June 2012.

[4] Shouyi Wang, Cheng-Jhe Lin, Changxu Wu, and Wanpracha Art Chaovalitwongse "Early Detection of Numerical Typing Errors Using Data Mining Techniques" IEEE November 2011.

[5] Mahdi Esmaeili, Amirhosein Mosavi, "Variable Reduction for Multi Objective Optimization Using Data Mining Techniques; Application to Aerospace Structures" 2010 2nd International Conference on Computer Engineering and Technology.

[6] Sérgio Ramos, Zita Vale "Data Mining techniques application in Power Distribution utilities" IEEE 2008.

[7] Dominik Fisch, Edgar Kalkowski, and Bernhard Sick "Knowledge Fusion for Probabilistic Generative Classifiers with Data Mining Applications"IEEE Transactions (Vol. 26) 3, March 2014.

[8] Jianlin Xu, Yifan Yu, Zhen Chen_, Bin Cao, Wenyu Dong, Yu Guo, and Junwei Cao "MobSafe: Cloud Computing Based Forensic Analysis for Massive Mobile Applications Using Data Mining".

[9] Tsinghua Science and Technology ISSN (Vol. 18), 4, August 2013, 418–427.

[10] P.J. Sandford"Detecting security threats in network core using data mining techniques" ISSN-1-4244-0143-7/06/20 IEEE 2006.

[11] Maral Haghighat, Hamid Rastegariand, Nasim Nourafza "A Review of Data Mining Techniques for Result Prediction in Sports" ACSIJ Advances in Computer Science: an International Journal, Vol. 2, Issue 5, No. 6, November 2013. ISSN: 2322–5157.

[12] Elovici, Y., Kandel, A., Last, M., Shapira, B., Zaafrany, O., "Using Data Mining Techniques for Detecting Terror-Related Activities on the Web".

[13] Romero, C., Ventura, S. and De-Bra, P. "Knowledge Discovery with Genetic Programming" for Providing Feedback to Courseware Authors, Kluwer Academic Publishers, Printed in the Netherlands, 2004".

[14] Dongsong Zhang and Lina Zhou " Discovering Golden Nuggets: Data Mining in Financial Application" IEEE Transactions, Vol. 34, 2004.

[15] Dongsong Zhang and Lina Zhou " Discovering Golden Nuggets: Data Mining in Financial Application" IEEE Transactions Vol. 34, 2004.

[16] Neelamadhab Padhy, Dr. Pragnyaban Mishra, Rasmita Panigrahi " Survey of Data Mining Applications And Feature Scope" (IJCSEIT), Vol. 2, No. 3, June 2012.

Intrusion Detection System for Wireless Networks: A Review

[1]Vikas Singla, [2]Monika Sachdeva and [3]Sunil Kumar Gupta
[1]Punjab Technical University, Jalandhar
[2]Shaheed Bhagat Singh State Technical Campus, Ferozpur
[3]Beant College of Engineering & Technology, Gurdaspur
E-mail: [1]singla_vikas123@yahoo.com, [2]monika.sal@rediffmail.com, [3]skgbcetgsp@gmail.com

***Abstract*—Wireless networks have recently been gaining widespread deployment and they can be easily attacked as compared to wired networks. CERT statistics reports state that the amount of intrusions on network has excessively increased year by year. With the increasing amount of network throughput and security threat, the study of intrusion detection systems (IDSs) has received a lot of attention throughout the computer science field. The traditional wired IDS is a great system, but unfortunately it does little for the wireless world. In this paper we will do the comprehensive survey and review the current security techniques and some IDS systems for protection against various types of attacks in wireless networks.**

Keywords: Wireless Network, Intrusion Detection, Attacks, Security

I. Introduction

In today’s arena, Wireless Networks are more popular and better alternative as compared to wired networks. In today’s life, we will find wireless networks everywhere e.g. at homes, an offices, or at business places. The development of wireless networks offers the promise of a flexible, low cost solution for monitoring critical infrastructure. The biggest concern with wireless network has been security. Security methods are designed in order to avoid unauthorized access to system assets and information. Be that as it may, totally avoiding breaks of security, at present, doubtful. However, We can attempt to identify these Intrusion endeavors. This field of research is known as Intrusion Detection. Intrusion detection system is method which secure our network from various kind of attacks [1, 4]. The main purpose of intruder is to hack the important information in the network, like using the bandwidth of node or increasing the delay time in providing the services over the network under consideration.

Other Sections in the paper are structured as follows. Section 2 of the paper defines the various security goals. Section 3 gives classification of different attacks on wireless network. Section 4 presents the Literature Review. Finally, we are concluding the paper with some goals for future work.

II. Security Goals

Any routing protocol must have an essential set of security mechanisms. These type of mechanisms help to prevent, detect, and respond to different types of security attacks [6]. Different types of security goals are required to be dealt with for maintaining a reliable and secure ad-hoc network. These are as below:

1. *Authentication*: This is concerned with unintended users are not be authorized to enter into the network. Authentication is assurance that the user tries to enter into the network is authentic user what it is claiming and it is not the intruder. An attacker tries to impersonate the user and thus getting unauthorized access to network and sensitive information of the network. Authentication is not to allow such types of users to access the network.
2. *Confidentiality*: Protection of any information from being exposed to unintended users. This is concerned with sending the message in such a way that unintended users cannot read the actual contents of the message. In ad-hoc networks confidentiality is more difficult to achieve as in the ad-hoc networks intermediates nodes receive the packets from other recipients.
3. *Availability:* Services are required to be provided when required. Availability is basically concerns with the availability of network. It has no concern with the actual data sent over the network. On the physical and media access control layer attacker can use jamming techniques to interfere with communication on physical channel. On network layer the attacker can disrupt the routing protocol.
4. *Integrity*: Message being transmitted is never altered. Data integrity has more importance than confidentiality. The reason is that by means of confidentiality, the attacker is only able to see the physical environment. And by means of reading the data, they are only able to read where the sensors are actually placed..But if the attacker is successful in alteration of data, the deployer of wireless sensor network will not be able to know what is actually happening in the network.

III. Classification of Attacks

Due to their underlined architecture, wireless networks are more easily attacked than a wired network. The attacks on wireless routing protocols can be broadly classified into following categories [4, 5, 6].

1. *Passive Attacks*, has no concern with the operation of the protocol. Passive attacks try to find the useful information about the network. Suck kind of attacks tries to collect the routing information

by sniffing the network. Passive attacks are difficult to detect. Due to this providing security against passive attacks is more complex. Using passive attacks, attacker can identify the network topology, however using these attacks it is difficult to find the actual location of node.

2. *Active Attacks*, the main function of these type of attacks is to disrupt the operation of the protocol by inserting arbitrary packets into the network. This kind of attacks tries to get authentication into the network so that they can access all the packets travelling over the network and disabling the operation of network.
3. *Denial-of-Service attacks (DoS)*, cause a network to slow down or become unusable. In case of DoS attack, attacker generates traffic, insert this into the network, which in turn block the server for a long. Distributed Denial-of-Service attacks (DDoS) occurs when many computers are used against the target.
4. *Man-in-the-middle attack*, will occur when the attacker is able to read and edit the communications between the two parties without the parties are being aware of the presence of attacker.

IV. Literature Review

The various solutions provided by different authors from above said attacks by providing some intrusion detection system are discussed below.

Madhavi *et al.* [8] inspect the vulnerabilities of wireless network and contended that intrusion detection must be incorporated in the security system. They proposed a Mobile Intrusion Detection System suitable for wireless networks, which distinguishes nodes misbehavior, irregularities in packet sending, for example, some nodes dropping packets. Proposed System does depend on overhearing packet transmissions of neighboring nodes. Proposed System sets the various thresholds dynamically.

Biradar *et al.* [10] proposed a security based multicast routing mechanism in MANET. Proposed method finds multicast routes to receivers by calculating route request packets and route reply packets. Performance of the proposed method is compared with on-demand multicast routing protocol and enhanced on-demand multicast routing protocol. They presumed that the proposed method delivers better PDF, reduced packet delay and reduced overheads.

Bhatnagar *et al.* [11] examined about issues and difficulties of IDS system for wireless sensor network and suggested a secure method that can recognize possible intrusion in the network, alarming client after intrusion had been discovered and reconfigure the system. In this paper, authors are mainly focused in multi hop WSNs and proposed an intrusion detection system using decision making technique.

Table 1 Literature Survey

Author	Publication Year	Proposed IDS Scheme/ Technology	Work Done	Conclusion
Madhavi *et al.,* [8]	2008	Mobile Intrusion Detection System	Propose an MIDS Suitable for Multi-Hop Ad-Hoc Wireless Networks, Which Find Out Misbehavior Nodes and Packet Forwarding Anomalies.	They Proposed MIDS Which Detects Packet Drops or Delays that Violate the Respective Flow Requirements.
Biradar *et al.* [10]	2010	A stability based multicast routing scheme	Performance of the method suggested by them is compared with ODMRP protocol (on-demand multicast routing protocol) and EODMRP protocol (enhanced on-demand multicast routing protocol).	Method suggested by them provided good PDR (packet delivery ratio), less packet delay and less overheads.
Bhatnagar *et al.* [11]	2010	Decision Making Technique	Discussed about various challenges in intrusion detection system for wireless network and proposed a new method for securing the network.	Proposed intrusion detection system defenses the strength of a wireless sensor networks using decision making technique.
Ming-Yang Su [12]	2011	Anti-Blackhole Mechanism	Detect and separate malicious nodes.	When the suspicious value exceeds the threshold value, an IDS nearby will send a broadcast message to all nodes saying them to cooperatively isolate the malicious node.
Sharma *et al.* [13]	2011	Misuse Detection System	Proposed a new Network Intrusion System that detects the Denial of Service(DoS) attack of Wireless Network.	The proposed method provided the safer transmission in Denial of Service and Man in Middle Attack.
Mulert *et al.* [14]	2012	Reactive intrusion detection node blacklisting scheme.	Analysis of SAODV to identify unresolved threats to the algorithm, such as medium access control layer misbehavior, resources depletion, black holes, worm- holes, jellyfish and rushing attacks.	Provide solution to various threats in MANET using AODV and SAODV

Ming-Yang Su [12] provided a mechanism for finding and separating the malicious nodes in the network. All IDS nodes perform a mechanism known as Anti-Black-hole mechanism, which assesses the suspicious estimation of a node by calculating difference between RREPs and RREQs transmitted over the node. At the point when a suspicious value exceeds the threshold value, an IDS adjacent will broadcast a block message, advising all nodes on the system, requesting them to helpfully disconnect the malicious node.

Sharma *et al.* [13] proposed an Network Intrusion System that will detect the Denial of Service Attack. The proposed method will finds the intrusion, on the bases of the Misuse Detection which has less false negative. Proposed System detects the intruders by the IP address.

Muler *et al.* [14] worked on networks using AODV and Secure AODV Protocols. They conducted a vulnerability analysis of SAODV to recognize uncertain threats to the algorithm, for example, medium access control layer misconduct, assets consumption, black holes etc. They contrast this helplessness investigation and proposed method to handle the distinguished attacks. They proposed method that incorporate multipath routing, incentive schemes, directional antennae, packet leashes etc.

V. Conclusion and Future Scope

In this paper, we present a review of recent work on different approaches of Intrusion detection system for wireless networks. Each technique has its own superiority and limitations, so that we should be cautious about selecting the technique. We provide a table which summarized the work of different authors to easily grasp the overall picture. We provide a comprehensive review of IDSs. However, there remain many open issues and future challenges.

Acknowledgment

Authors wish to thank Punjab Technical University, Jalandhar for providing inspiring environment for research work.

References

[1] Cabrera, J.B.D., Ravichandran, B & Mehra R.K., "Statistical Traffic Modelling for Network Intrusion Detection", In Proceeding of the IEEE Conference (2000).

[2] Y. Zhang, W. Lee., "Intrusion detection in wireless ad-hoc networks", In Mobile Computing and Networking, (2000), pp. 275–283.

[3] C. Tseng, P. Balasubramanyam, C. Ko, R. Limprasittiporn, J. Rowe and K. Levitt, "A specification-based intrusion detection system for AODV." In Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, (2003), pp. 125–134.

[4] W. Zhang, R. Rao, G. Cao and G. Kesidis, "Secure routing in ad hoc networks and a related intrusion detection problem." In Proceedings of the IEEE Military Communications Conference, (2003), pp. 735–740.

[5] A. Mishra, K. Nadkarni, A. Patcha, "Intrusion Detection in Wireless Ad-Hoc Networks." In IEEE Wireless Communications, (2004), pp. 48–60.

[6] A. Perrig, J. Stankovic and D. Wagner, "Security in Wireless Sensor Networks", In Communications of the ACM, Vol. 47, No. 6, (2004), pp. 53–57.

[7] S. Kurosawa, H. Nakayama, N. Kato, A. Jamalipour, Yoshiaki, Nemoto, "Detecting black hole attack on AODV-based Mobile Ad Hoc Networks by dynamic learning method", International Journal of Network Security 5 (2007), pp. 338–346.

[8] S. Madhavi, Tai Hoon Kim,"An Intrusion detection system in mobile adhoc networks", International Journal of Security and Its Applications, Vol. 2, No. 3, (2008) pp. 1–16.

[9] R. Huang, Y. Zhuang, Q. Cao, "Simulation and Analysis of Protocols in Ad Hoc Network", International Conference on Electronic Computer Technology IEEE (2009).

[10] R. Biradar, S. Manvi, M. Reddy, "Link stability based multicast routing scheme in MANET", Computer Networks 54 of Elsevier (2010), pp. 1183–1196.

[11] R. Bhatnagar, A.K. Srivastava, A. Sharma, " An Implementation Approach for Intrusion Detection System in Wireless sensor Network", International Journal on Computer Science and Engineering Vol. 02, No. 07, (2010), pp. 2453–2456.

[12] M.Y. Su, "Prevention of selective black hole attacks on mobile ad hoc networks through intrusion detection systems", Computer Communications 34 (2011) pp. 107–117.

[13] M. Sharma, Anuradha, " Network Intrusion Detection System for Denial of Service Attack based on Misuse Detection", International Journal of Computational Engineering & Management Vol. 12, (2011) ISSN (Online): 2230–7893, pp. 19–23.

[14] J.V. Mulert, I. Welch, W.K.G Seah, "Security threats and solutions in MANETs: A case study using AODV and SAODV", A Journal of Network and Computer Applications 35 (2012) of Elsevier, pp. 1249–1259.

Taxonomy of Network Layer Multicast Routing Protocols in Mobile Ad-hoc Networks

Kanwalpreet Kaur[1], Krishan Kumar Saluja[2] and Rajdeep Singh[3]
[1,3]Department of Computer Science & Engineering,
Punjab Institute of Technology (PTU Main Campus), Kapurthala, India
[2]Department of Computer Sc. & Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab, India
E-mail: [1]kanwal.preet2007@gmail.com, [2]k.salujasbs@gmail.com, [3]ptuap_cse@yahoo.com

***Abstract*—A MANET consists of self organizing mobile nodes which exhibit dynamic behavior during their multicast operations. Thus, it is imperative to obtain the best way to provide multicast services in this kind of environment. For this purpose investigation and quantification of existing multicast routing protocols is the foremost step. In this article detailed discussion is done regarding basic behaviors of the multicast protocols and the types of services provided by them. Multicast protocols have different layers of operation namely, network layer, application layer and MAC layer. This work presents the coherent survey of existing network layer multicast routing protocols and discusses their routing mechanisms and the application/services. The classification of protocols on the basis of their types of routing mechanism and type of application/services, provide the comprehensive information about the protocols. Thus this paper aims to present a clear view to the MANET researchers and application developers so that they could select the multicast protocol accordingly for their work.**

Keywords*—*MANET, Multicast, Taxonomy

I. Introduction

A MANET is a self organized network consisting of mobile nodes. The nodes act as the both host and the router. MANETs have many constraints like channel efficiency, power related problems, security, packet drops and noise errors. So if there are more no of receivers than it is better to use a multicast routing protocol rather than unicasting the data from source to the reciever. A multicast routing protocol sends multiple copies of datagram to the intended receivers. Thus with the help of multicast routing protocols energy consumption, routing and processing delay and cost of communication gets reduced [1].

Apparently, many classification criteria have been proposed for multicast routing protocols. In this work classification criteria have been chosen in such a manner that most of the common mechanisms employed by the well known multicast protocols are covered. The details of the application or services provided by them are also discussed simultaneously. Most of the survey papers classify multicast protocols on the basis of multicast topology or initialization approach only. This paper presents the state-of-the-art review for multicast protocols operating at the network layer by introducing new technical trends and the avenues of the research examples being carried out in this field.

Primary goal of this survey is to provide precise and an up to date useful taxonomy. To achieve this goal the basic properties of the multicast protocols of the network layer are first identified and multicast protocols are then classified according to the routing mechanisms and types of services they provide. Already existing multicast protocol designs are then summarized based on the proposed classification criteria and are referenced for future investigations. As compared to the previous surveys, this research provides the wider view of the different operational features of multicast routing protocols for MANETs. The key contributions of this research are as follows:

1. To classify network layer multicast protocols according to the type of routing mechanism and type of application/service delivered by the protocol simultaneously, to help researchers to analysis and compare the network layer protocols more easily.
2. To granulize routing mechanisms to deeper level, to study the protocols more minutely.
3. To identify and distinguish the main applications/services provided by multicast protocols.
4. To review typical multicast routing protocols according to the proposed classification criteria.

This paper is further organized as follows. In Section II, taxonomy of multicast routing protocols is presented. Further in Section III, comprehensive survey of typical multicast routing protocols based on the proposed taxonomy is discussed. Then later on, in Section IV research work is concluded.

II. Taxonomy of Multicast Routing Protocols

Apart from the fact that each multicast routing protocol has some distinct characteristics of its own, they exhibit some common features too, on the basis of which they can be categorized and studied readily. So in this work we have classified the protocols primarily according to their routing mechanism types and the type of service/application they provide as shown in Fig. 1. These two chief categories provides the basic features of a multicast protocol, concurrently specifying the main services provided by that protocol.

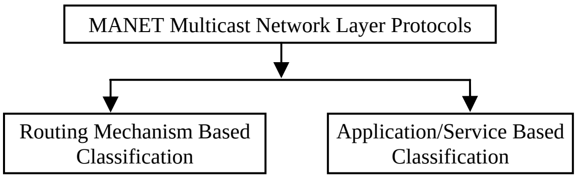

MANET Multicast Network Layer Protocols
Routing Mechanism Based Classification
Application/Service Based Classification

Fig. 1 Principal Classification Criteria

A. *Routing Mechanism Based Classification[2]*

This classification is basically concerned with the various features exhibited by multicast protocols during their routing operation. The protocols are categorized on the basis of type of network maintenance approach, multicast topology used for multicasting, routing scheme of the protocol, multicast initialization approach followed, type of core mechanism and dependency on any underlying unicast protocol. Fig. 2 presents the overview of this classification. So it covers nearly all the features which a multicast routing protocol can have.

Moreover each category is further divided into sub categories to give the exact and precise information of that particular routing mechanism. So this classification covers a huge no. of multicast protocols providing the inclusive details and discussions of their operational features. The main categories under this classification and the analogous subcategories are explained as follows:

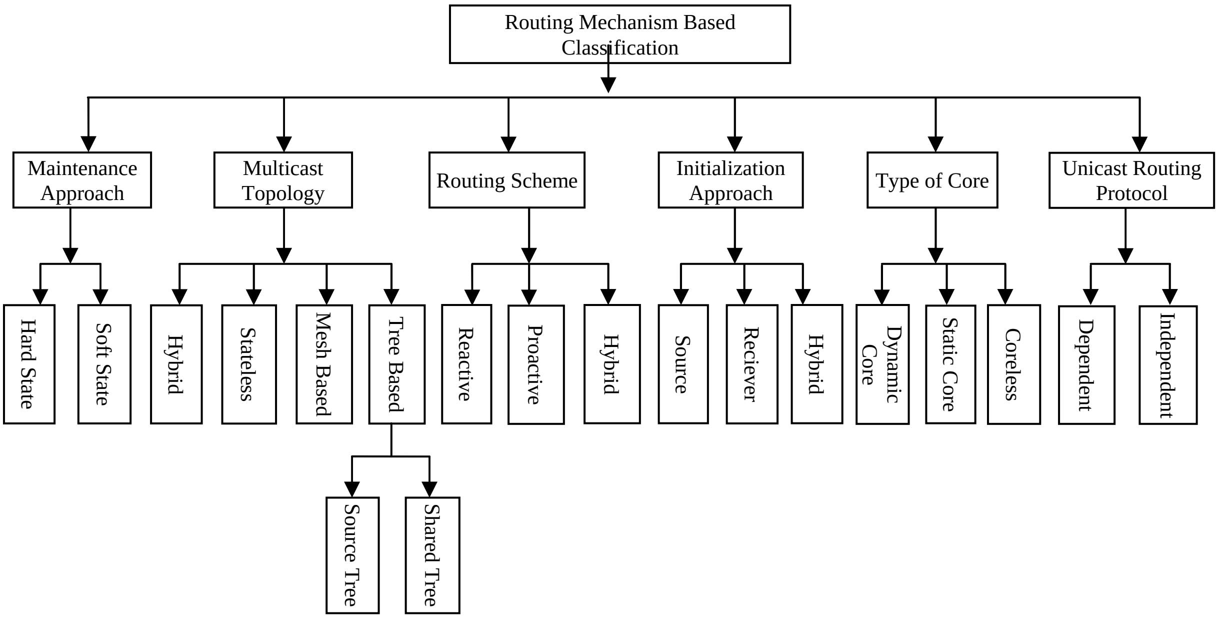


Fig. 2 Routing Mechanism Based Classification

1) *Maintainence Approach [2]*

As we discussed earlier that MANETs have dynamic environment and frequent topology changes occur due to the frequent link breakages. So updating routing tables and information of the participating nodes promptly, becomes a necessity to keep consistency of multicast routing topology. Maintenance mechanisms can be chiefly of two types–soft state based and hard state based. The details of these two approaches are as follows:

1. Soft-State Approach:In this approach the control packets are flooded periodically to obtain the multicast group membership updates. The state of the connection between the nodes is checked periodically and multicast group information is refreshed. This approach is flexible for use in dynamic wireless communication and provides reliability.
2. Hard-State Approach:In this approach, the information of broken links is delivered by two methods. The implicit control packets are sent reactively in the first method when a link breakage occurs. In the second proactive method the link breakages are predicted through GPS or local prediction mechanisms and routes are repaired or updated accordingly beforehand. The successful implication of this approach involves many crucial factors like on time notification of failure, quick initialization of the repair process and a speedy link repair mechanism. It involves less overhead as compared to soft state approach.

2) *Multicast Topology [2]*

It is the most common criteria for the classification of multicast protocols. It classifies protocols on the basis of how routes are constructed and how mobile nodes arrange themselves to do the multicast operation. So according to this classification criteria, multicast protocols can be classified into four types – mesh based, tree based, hybrid or stateless. The tree based protocols can be further sub categorized according to the multicast tree formation and operation i.e. source tree based or shared tree based multicast protocol. These types can be further explained in detail as follows:

1. *Tree based Topology:* In this topology, a single and a shortest path exists between a source and a

destination node. It can be further classified into two types:

- *Source Tree based:* A different multicast tree is built for each source node. So a source must have the information of the addresses of the receivers.
- *Shared Tree based:* In this topology, a single multicast tree is created for all source nodes, rooted at a node called as a core node, responsible for the overall management of the topology and multicast group. Shared tree approach is less efficient than source tree approach because the paths constructed between the source and destination node are not the shortest. Moreover it has to keep more routing information so more control overhead is associated with it. Due to the presence of core node, there exists the single point failure threat.

2. *Mesh Based Topology:* In this topology, the packets are forwarded on the set of interconnected nodes forming a mesh structure called as the forwarding group. Route establishments are either done by forwarding the control packets or with the help of core nodes. So redundant paths are available from source to destination and thus give more stability in case of mobile scenarios. Unlike tree based topologies, reconstruction of topology is not required in case of mobile nodes. As there exist a route between a source and destination always due to mesh structure. So provides high robustness but is less efficient than tree based approaches.
3. *Hybrid Topology:* It combines both the features of tree and mesh topologies i.e. robust and efficient. But they can produce non-optimal trees with nodes having mobility so efficient mechanisms for managing group membership information and nodes mobility are required.
4. *Stateless:* Both the tree based and mesh based approach involves huge control overhead. So, stateless topologies are used to minimize control overhead. In this routing information of all forwarding nodes is not maintained rather, source node explicitly specifies the destination nodes list and data is directly sent to those nodes making it suitable for a small multicast group.

3) *Routing Scheme[3]*

There are principally three ways to update the routing information among the mobile nodes in case of MANETs. So protocols can be classified in following three types on the basis of routing schemes or approaches they follow:

1. *Proactive:* This approach is also known as the table driven approach because each node has a table representing the topology of the network. To update this information in the tables the topology information is exchanged between the nodes from time to time. So in this routing scheme, the information about network is maintained at each node irrespective of the fact that whether the information is needed or not. So this leads to more power consumption and more control overhead.
2. *Reactive:* This approach is also known as On Demand approach. The routes are created only when desired by source node. The group membership information is updated on demand. This routing scheme is more scalable than proactive scheme and does not need maintenance of whole network information thus requires less control packets. Path discovery process is more difficult in this case.
3. *Hybrid:* This combines the features of above two approaches to alleviate the problems in them. Zones are maintained and different routing scheme are deployed at different zones.

4) *Initialization Approach [4]*

A multicast operation is initiated by a single node. It can be a source node or a receiver node. So on the initialization approach basis, the multicast protocols can be classified into following types:

1. *Source based Initialization [5]:* In this method, the source node offers the data to the interested set of receivers. So this method is sometimes called as pushing. The receiver nodes acknowledge each packet sent by the source node. Thus source takes the responsibility for data delivery and processes feedback from the receivers. This method is more suitable for the dense groups i.e. when no. of receivers are higher. But when no of sources or senders increase, control overhead too increases exponentially.
2. *Receiver based initialization:* In this method, the reciever looks up for the senders of the desired data. So this method is also known as pulling. In this case, it is the responsibility of receiver to detect transmission error and packet loss by checking the gaps in the sequence no. of received data packets. This method is well-suited for sparse groups i.e. when no. of receivers is lesser. When no. of senders increase, the control overhead too increases but linearly.
3. *Hybrid [2]:* Some protocols do not fall strictly in any category. The initialization is sometimes done by source or by the receiver; this is called as the hybrid approach.

5) *Type of Core[4]*

Two approaches can be used in a multicast group to give the network information: distributed or centralized. The nodes which maintain this network information and do the membership management are called as the core nodes. So, multicast protocols can be classified on the basis of types of core as follows:

1. *Coreless:* In case of distributed approach there is no particular node which has the complete membership information so this approach is called coreless. But large control overhead is associated with distributed approach, because each node keeps the information and exchanges with the neighbours.
2. *Static Core:* In the centralized approach, the membership information is maintained by a single node called as the core node. When the core node is assigned by an external entity before the multicast session establishment, then it is called as the static core approach. The major drawback of this approach is that if the core node fails, the whole membership information will be lost and multicast group will collapse.
3. *Dynamic Core*: In this approach, the core node is selected dynamically. So if the present core node fails, the new core is selected by the members dynamically. Core based approaches have relatively less control overhead because control packets are sent to the core and then from the core to the members. Moreover, dynamic core approach is more stable in high mobility.

6) *Unicast Routing Protocol [4]*

Many multicast routing protocols work on some underlying unicast routing protocols while others work independently. So in this context classification can be done in two ways :

1. *Independent multicast protocols*: These protocols do not need a unicast protocol for their operation. They have inbuilt unicast routing protocol and are designed in the manner to support both multicast and unicast simultaneously.
2. *Dependent multicast protocols:* They can be further divided into two subparts. Some multicast protocols can work only with specific unicast routing protocols while some can work with any available unicast routing protocol. It suffers from higher control overhead.

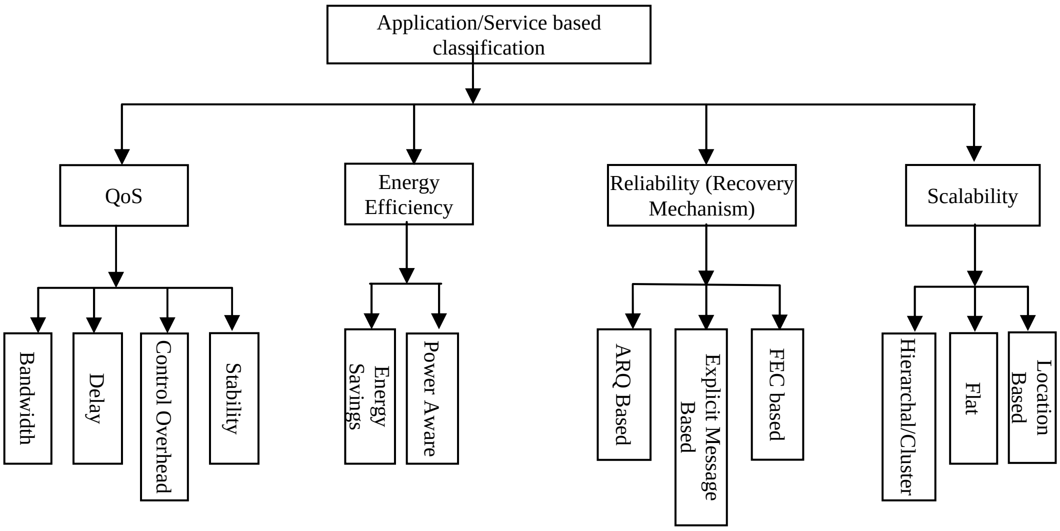


Fig. 3 Application/Service Based Classification

B. *Application /Service Based Classification[6]*

Each multicast protocol is designed for a specific purpose. So different multicast protocols provide different services and are used in different applications. The protocols are categorized on the basis of type of QoS they provide, whether the focus is on energy savings or protocol is a power aware protocol. Other classification criteria include reliability as a service, which is further categorized on the basis of recovery mechanisms being used and the last classification criteria is scalability. Fig. 3 represents the various classification criteria on the basis of application/service. They can be further explained as follows:

1) *Type of QoS*

QoS is one of the important elements to evaluate the performance of MANETs because QoS constraints the bounds on delay, jitter, bandwidth and control overhead. QoS routing provides route from source to destination and end to end QoS simultaneously. It further provides the reliable communication by providing the stability in terms of route, link or node stability. But in case of MANETs, it is difficult to provide QoS because of sharing of bandwidth and mobility of nodes. Protocols can be classified on the basis of type of QoS they provide and can be divided chiefly into below four categories. Most of the network layer multicast protocols can be categorized into these subtypes:

1. *Bandwidth:* It is the amount of data transferred per second. Some multicast protocols are designed to utilize the bandwidth properly because real time applications require assured bandwidth for standard and continuous presentation of data. So these protocols provide the better utilization of bandwidth.
2. *Delay:* It is the time taken by the data packet from source to destination. Most of the multicast protocols are designed for multimedia applications where delay boundsare quite critical. So some protocols are designed particularly to keep the end to end delay least. End to end delay includes– packet compression and packetization, transmission, queuing and synchronization, decompression and depacketization at the destination end. So by focusing on delay, these protocols prove to be as a boon for multimedia applications.
3. *Control Overhead:* It is the total control packets used per data packet delivered. If the control overhead will be higher; more bandwidth will be utilized for control packets than data. The protocols which eradicate this problem are included in this category.
4. *Stability:* Stability is mainly in the context of – node, link and route. A node's stability relies on the mobility, life of battery, no. of interfaces currently being used and the data transmission rate. Higher the mobility, lesser is the stability. More is the battery life, higher will be the stability. If no. of interface is more, then energy spent will also be more, resulting in lesser stability.

2) *Energy Efficiency[5]*

In MANETs nodes have limited energy supply and due to adverse network conditions in MANETs, it becomes difficult to save the energy of batteries. So designing of energy efficient protocols is one of the major issues and many multicast protocols are designed only for this purpose. They can be further classified into following categories:

1. *Energy Savings:* In this approach, the primary goal of protocol is to find a routing path with least energy consumption.
2. *Power Aware:* The primary focus is to consume node energy in a balanced manner using a cost function and keeping track of the node's residual battery capacity. Thus all the links of the nodes and the power consumption of the nodes can be reduced by managing the transmission power of node wisely.

3) *Reliabililty[4]*

A protocol is considered as a reliable if it has mechanisms for error detection and to indicate the source or destination by sending error messages and availability to retransmit the lost packets again. So, due to frequent link changes in the MANETs it becomes a very challenging task to provide reliability. On the basis of various recovery mechanisms used, the multicast protocols can be divided into following main categories:

1. *ARQ (Automatic Retransmission Request) based:* They are called as the deterministic protocols. They are further of two types- sender initiated and reciever initiated. in case of sender initiated protocol the ACK messages are used and sent back by the receivers for the retransmission of the data packets. NACK messages are sent in receiver initiated based protocols after detecting missing packets.
2. *Explicit Message (Gossip) based:* In this an explicit message, sometimes called as a gossip message is transferred in a peer to peer manner. It consists of the information about the multicast packets received and missing packets. They do not guarantee full delivery of packets.
3. *FEC (Forward error Correction) based :* In this the reliability is provided by repeatedly sending the data. The data is encoded and then split into fragments. The receiver receives the fragments and reassembles the packets. But this approach is more suitable in the scenario where loss rates are predictable which is difficult in MANETs.

4) *Scalability*

The multicast protocols are scalable with respect to some constraints posed by the MANETs. They can be further categorized into following three types:

1. *Flat:* The homogeneous nodes in terms of network resources and computing power constitute flat network architecture. The protocols having flat network architectures are included in this category.
2. *Hierarchical:* These protocols have physically hierarchical architecture. The multicast structures are built at each level of hierarchy for efficient multicast delivery.
3. *Location Based:* In these protocols, the availability of a Global Positioning System (GPS), Bluetooth or other location systems is required to get the geographical information of the multicast networks. The sender determines the location of the destination by using the location service. Moreover the routing decisions of each forwarding node relies on it's neighbours and destination node.

III. Comprehensive Survey of Typical Multicast Routing Protocols

Multicasting efficiently supports many applications. The multicast protocols are driven by specific goals and needs based on suppositions about the network or application. Each protocol has its own pros and cons. It is difficult to cover all the multicast protocols proposed so far, in a single review. In this section the survey of recent and popular multicast protocols which operate on the network layer is done according to the proposed classification criteria. The Table 6 below lists the various network layer multicast protocols and categorizes them according to the classification criteria proposed.

IV. Conclusion

The purpose of this paper is to propose the classification criteria on the basis of different routing mechanisms and the application/services provided by the multicast protocols simultaneously. The categorization of the routing selection principles can simplify the work of a network designer. This paper aims to provide a useful survey to the researchers or the beginners who are going to embark on MANETs. It can be concluded that each protocol satisfies the maximum possible requirements but one size does not “fit all”. To design a multicast protocol which meets all the requirements, is a very complicated task and will be difficult to operate in MANETs environment.

Table 1 Comprehensive Survey of Typical Multicast Routing Protocols

	Classification Criterion									
			Routing Mechanism based Classification						***Application/Service based Classification***	
	S.No	***Name***	***T***	***M***	***R***	***I***	***C***	***UD***	***Type***	***Sub Type***
List of Protocols	1.	ABAM[2]	SOT	HS	RE	SRC	CL	IND	QoS	ST
	2.	ABMRS[2]	ME	HS	RE	SRC	CL	IND	REL	EM
	3.	ACMRP[2]	ME	SS	RE	SRC	DYC	IND	QoS	ST
	4.	AMRIS[2]	SHT	HS	RE	SRC	CL	IND	REL	ARQ
	5.	DDM[5]	S	SS	RE	REC	STC	IND	QoS	BW
	6.	DQMRP[3]	SHT	SS	PRO	SRC	CL	IND	QoS	D
	7.	EHMRP[3]	THY	SS	RE	SRC	DYC	IND	SC	HL
	8.	EODMRP[3]	ME	SS	RE	SRC	CL	IND	REL	EM
	9.	FGMP[2]	ME	SS	RE	REC	CL	DEP	QoS	CO
	10.	HQMRP[3]	S	SS	PRO	REC	CL	IND	QoS	BW
	11.	HZMAODV[6]	SHT	HS	RE	REC	DYC	DEP	SC	LO
	12.	LAM[3]	SHT	SS	RE	SRC	STC	DEP	SC	FL
	13.	LSMRM[6]	ME	HS	RE	SRC	CL	IND	QoS	ST
	14.	MAMR[5]	THY	HS	RE	IHY	CL	DEP	QoS	D
	15.	MAODV[2]	SHT	HS	RE	REC	DYC	DEP	QoS	ST
	16.	MCEDAR[2]	THY	HS	PRO	IHY	DYC	DEP	QoS	BW
	17.	MMAs[2]	SHT	HS	RE	REC	DYC	DEP	EGY	ES
	18.	NSMP[2]	ME	SS	RE	SRC	CL	IND	QoS	CO
	19.	ODMRP[2]	ME	SS	RE	SRC	CL	IND	QoS	ST
	20.	OGHAM[2]	THY	HS	RE	SRC	CL	IND	QoS	BW
	21.	OPHMR[2]	ME	HS	RHY	REC	CL	DEP	EGY	PA
	22.	PPMA[2]	SOT	SS	RE	SRC	CL	IND	QoS	ST
	23.	P-REMiT[5]	SOT	SS	PRO	SRC	CL	IND	EGY	ES
	24.	QARBE[6]	SOT	SS	RE	REC	CL	IND	QoS	BW
	25.	RDG[4]	S	SS	RE	SRC	CL	DEP	REL	EM
	26.	RMDP[2]	THY	SS	RE	SRC	CL	IND	REL	FEC
	27.	SPBM[3]	SOT	SS	RE	REC	CL	IND	SC	LO
	28.	SRMAODV[6]	SHT	HS	RE	REC	DYC	DEP	REL	EM
	29.	SRMP[2]	ME	HS	RE	REC	CL	DEP	EGY	PA
	30.	WBM[2]	SOT	HS	RE	REC	CL	IND	QoS	BW

T-Multicast Topology M-Maintenance Approach R-Routing Scheme I-Initialization Approach C-Type of Core UD-Dependency on Unicast Routing Protocol THY-Hybrid Topology S-Stateless Topology ME-Mesh Topology T-Source Tree based topology SHT-Shared Tree based Topology Hard State Maintenance SS-Soft State Maintenance RHY-Hybrid Routing RE-Reactive Routing PRO-Proactive Routing IHY-Hybrid Initialization REC-Reciever initialized SRC-Source based initialized DYC-Dynamic Core STC-Static Core CL-Coreless DEP-Dependent on unicast protocol IND-Independent of unicast routing protocol QoS-Quality of Service BW-Bandwidth D-Delay CO-Control Overhead EGY-Energy Efficiency ES-Energy Savings PA-Power Aware REL-Type of Reliability Mechanism ARQ-ARQ based reliability EM-Explicit Message based Reliability FEC-Forward Error Correction Based Reliability SC-Type of Scalability Approach HL-Hierarchical/ Cluster Scalability FL-Flat Scalability LO-Location based scalability ST-Stability

ACKNOWLEDGEMENT

The authors take the privilege in extending gratitude to their family members and friends who rendered their support throughout this work.

REFERENCES

[1] S. Corson and J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", RFC 2501, 1999.

[2] O.S. Badarneh, "Multicast Routing Protocols in Mobile Ad Hoc Networks: A Comparative Survey and Taxonomy", EURASIP Journal on Wireless Commun. And Networking, Vol. No. 26, 2009.

[3] L. Junhai, X. Liuand, Y. Danxia, "Research on multicast routing protocols for mobile ad-hoc networks", Computer Networks, Vol. 52, pp. 988–997, 2008.

[4] M. Ghasemi and M. Bag-Mohammady, "Classification of multicast routing protocols for Mobile Ad Hoc Networks" ICT Convergence (ICTC), Jeju Island, pp. 789–794, 2012.

[5] L. Junhai, Y. Danxia, X. Liu and F. Mingyu,"A survey of multicast routing protocols for mobile Ad-Hoc networks" IEEE Communications Surveys & Tutorials, Vol. 11, pp.78–91, 2009.

[6] R.C. Biradar, S. Kumar and S. Manvi, "Review of multicast routing mechanisms in mobile ad hoc networks", Journal of Network and Computer Applications, Vol. 35, pp. 221–239, 2012.

[7] G.U. Devi and R.S.D.W. Banu, "Performance Evaluation of Multicast Routing Protocols in Ad Hoc Networks", Parallel Distributed and Grid Computing (PDGC), India, pp. 757–761, 2012.

A Survey on LEACH and its Descendant Protocols in Wireless Sensor Network

Divya Prabha[1] and Vishal Kumar Arora[2]
[1,2]*Department of Computer Science & Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]divya.prabha431@gmail.com, [2]vishal.fzr@gmail.com

***Abstract*—Wireless sensor network is defined as wireless network of sensor nodes in which Routing technique is one of the most challenging issues. One of the major issues in WSNs is the limited battery power of the network sensor nodes. The battery power plays an important role in increasing the lifetime of the nodes. In WSN, routing among various routing technique, energy consumption is one of the most important consideration. To minimize energy consumption hierarchical routing protocols are the best known protocols. LEACH protocol is one of the most energy efficient clustering protocols. Leach increases the network lifespan by consuming a small percentage of the total dissipated energy in the network. We have surveyed the different hierarchical routing protocols, developed from the LEACH. The paper presents survey of LEACH protocol and its various descendant protocols like E-LEACH, TL-LEACH, M-LEACH, V-LEACH, LEACH-A, LEACH-B, LEACH-S. This paper also compares the above all routing protocols with original LEACH.**

Keywords: Clustering Leach Protocol, Network Routing, Wireless Sensor Network

I. Introduction

A WSN is a collection of wireless nodes with limited energy capabilities that may be mobile or stationary and located randomly on a dynamically changing environment which spread over a specific area [2]. Nodes in WSN are compact, light in weight and battery-powered devices. Therefore these special characteristics make all sensor nodes are allowed to communicate through a wireless medium and for the purpose of close-range sensing [5]. Routing in WSNs is very challenging from other wireless networks due to these specific characteristics [1]. These nodes are arranged randomly and can communicate among themselves to make an ad-hoc network. Sensor nodes are battery-powered and it is not easy to replace the batteries or recharge the batteries because each node has a limited energy supply. A clustered network is divided into various clusters. Within each cluster, one node is chosen as a cluster head (CH) among all sensor nodes and rest are treated as cluster members (CM). All sensor nodes co-operate each other to serve the requests. In each cluster, CH collects the data from the cluster members and relays the data either directly or via multi-hop transmission. Since the CHs utilizes more energy than the non-cluster heads. So, it distributes the workload of the CHs among the sensor nodes and their role is rotated among all nodes for energy-consumption equalization. One of the challenging issues in WSN is developing an energy-efficient routing protocol which increases the overall lifetime of the sensor network [6]. Hierarchical Routing is an efficient routing technique to reduce energy consumption LEACH (Low Energy Adaptive Clustering Hierarchy) is the first hierarchical routing protocol [6].

II. LEACH Protocol

LEACH is called “Low Energy Adaptive Clustering Hierarchy” is the first energy-conserving routing protocol and popular among all clustering algorithms for WSN. It reduces the energy significantly [1], [5]. In LEACH clusters are formed by distributed algorithm. First of all, a node is selected as a CH with a probability p and informs its decision to all nodes and after that each non-CH node determines its cluster by choosing the CH that can have least communication energy. The basic principle behind is that it assigns overall network’s energy consumption to each sensor node periodically. Thus, it can reduce the energy consumption and the lifespan of the entire network is prolonged. The role of CH is rotated within the clusters and among the nodes periodically in order to balance the load. This rotation is performed by each node by choosing a random number or threshold value T (n) between 0 and 1. If the random number < T (n), the node will become the cluster-head for the current round r, and other nodes join in the nearest cluster. After the completion of one period of data transmission, the network starts cluster reconstruction for the new round.

The threshold value is:

$$\begin{cases} P\ /1-p*(rmod1/p) \text{ if } nEG \\ T\ (n) = \\ 0 \text{ otherwise} \end{cases}$$

Where p is the desired percentage of CH nodes in the several sensors, r is the current round number, and G is the set of nodes which is not selected as a CHs in the last 1/p rounds [1], [4], [5], [6].

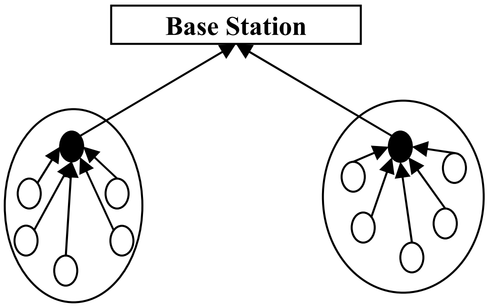


Fig. 1 Leach Clustering Hierarchical Model

III. Leach Protocol's Descendants

A. *E-LEACH-Enhanced-leach or Energy Low Energy Adaptive Clustering Hierarchy*

E-LEACH is evolved in order to solve the overload energy consumption problem and based on LEACH protocol [6].

This protocol has following some objectives:

- Cluster-head failure handling.
- To handle non-uniform and dynamic residual energy of the nodes.

The E-LEACH is based on the principle of same round concept as the original LEACH. E-LEACH is the enhancement of LEACH. It involves cluster head selection algorithms which have global information about the other sensors [4]. The total number of cluster-heads is a key factor which affects the overall performance of hierarchical routing protocols. If the number of CHs is less then each CH have to cover larger region, this will results the problem that some cluster-members get away from their CHs and hence it will consume more energy [6]. By considering the residual energy of sensor nodes as the main key factor, it decides whether that node should turn into the cluster head or not in the next-round. The communication between the cluster heads and the base station requires much more energy than common node; the larger number of cluster-heads will lead in increasing the energy consumption of the whole network and reduces the network lifetime. Therefore, it is necessary to choose optimal number of cluster heads for minimum energy consumption. E-LEACH uses the minimum spanning tree among cluster heads and choose that cluster head which has largest residual energy at the root node [5], [6].

B. *TL-LEACH (Two-Level LEACH)*

In original LEACH protocol, the Cluster Head (CH) collects and aggregates data from sensors within its own cluster and sends the information to the Base Station (BS) directly. Most of the time the Cluster Head(CH) can be located far away from the Base Station (BS), so it consumes most of its energy in sending information and then it will die faster in comparison of other nodes. Therefore, a next version of LEACH called Two-level Leach was evolved. In this leach protocol; Cluster Head (CH) collects data from other cluster members as original LEACH, but rather than relays data to the Base Station directly, it uses one of the Cluster Heads (CHs) that lies between the Cluster Head (CH) and the Base Station (BS) as a relay station [5]. The two-level structure i.e.TL-LEACH reduces the number of nodes which is used to transmit to the base station, and then effectively reduces the total energy consumption [3].

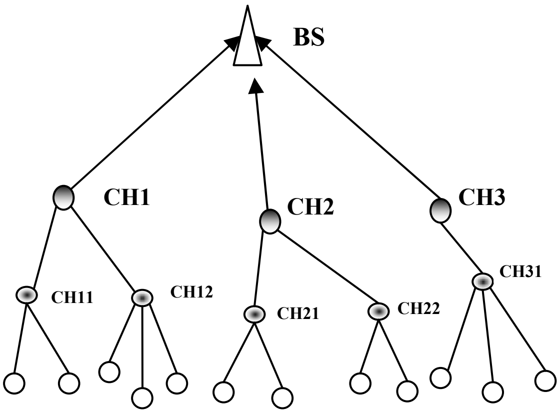


Fig. 2 TL-LEACH Protocol

C. *M-LEACH (Multi-Hop Leach)*

In LEACH the data is transferred from cluster head (CH) to base station (BS) node by using single hop communication and the distance between BS and CH has no effect [5]. The distance between the CH and the BS is increased when the network diameter is increased beyond a certain level [4]. Energy consumption will increase if distance is increased. Therefore in order to enhance the energy efficiency of the protocol this M-LEACH modifies original LEACH by allowing multi-hop communication used by sensor nodes within the cluster. This idea extends the existing solutions by using multi-hop communication in WSNs in which there is no direct communication between Cluster heads (CHs) or the sink due to the distance between them. Thus, the main idea of the solution is that the multi-hop approach is utilized inside the cluster and outside the cluster. Multi-hop Leach is a complete distributed clustering based routing protocol. Cluster Heads (CHs) also perform data fusion to the data receive by reducing the total transmitted and forwarded data in the network [4], [5], [6].

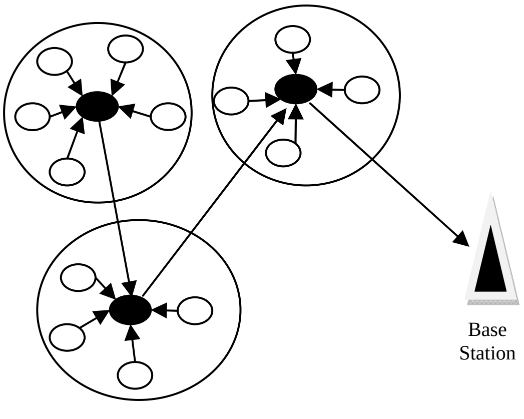


Fig. 3 M-LEACH Protocol

D. LEACH–C (Centralized LEACH)

The drawback to LEACH is that the number of CH nodes is little ambiguous to count [6]. To solve this problem LEACH-C has been proposed. LEACH-C is similar to original LEACH in operation except cluster formation [1]. In LEACH-C, centralized clustering algorithm involves. The steady state will remain the same but the setup phase of the LEACH-C is different. Each node broadcasts information about the current location and also the energy level to the base station. Thus base station utilizes the global information of the network and produce better clusters that requires less energy for data transmission [4]. The nodes with less than average energy are not consider in cluster heads selection whereas the nodes that have more than average energy, are selected for cluster heads. It uses GPS or the other location tracking method. Firstly the BS has to agree that only nodes with enough energy are allowed to participate in the selection of the CH and then after that base station (BS) broadcasts a message of the optimum cluster head IDs (Identifiers) to all nodes in the network. The node, having the same ID as the optimum cluster head ID, is nominated as a CH and transfer data by preparing TDMA schedule while remaining nodes wait for the TDMA schedule from their cluster heads [4], [6]. Leach-C gives deterministic threshold algorithm which considers the amount of energy in the node and/or whether or not the node was a CH recently. The number of CH nodes and its placement cannot give guarantee. Clusters can also be formed by using central control algorithm which may produce better clusters by distributing the CH nodes throughout the network [4]. Therefore, the selection criteria of cluster-head affect the performance and lifespan of the entire network.

E. V-LEACH (Vice-Cluster Head LEACH)

In the fundamental LEACH each cluster has a cluster head and when this cluster head does not have sufficient energy to transmit cluster member’s data to the base station, it dies. This is the main disadvantage of LEACH that when the cluster head dies all the data that is with it is lost. Another disadvantage of LEACH protocol is the random selection of cluster heads. There exists a probability that the cluster heads selected are unbalanced. They may remain in one part of the network and making some part of the network unreachable [6]. To overcome this problem V-LEACH has introduced the concept of alternate Cluster Head called Vice Cluster Head.

V-LEACH includes:

1. It is the responsibility of cluster head to transmit the data to the base station the it receives from the cluster members.
2. A vice-CH defined as that node which will become a CH of the cluster when the existing CH dies.
3. Cluster nodes used in gathering data from the environment and send the gathered data to the CH.

Therefore, in V-LEACH protocol, besides having a CH in the cluster, there is also a vice-CH which has replaced the role of the CH in case the CH dies because of the reasons mentioned above [5]. But it does not give a solution to the problem when Vice-CH Dies and then the network start dissipating the energy very quickly and finally the network dies completely. In V- LEACH the CH and Vice-CH are selected on the basis of Energy, Distance and Residual Energy. The V-LEACH will increase the overall network life and improve the total communication over the network.

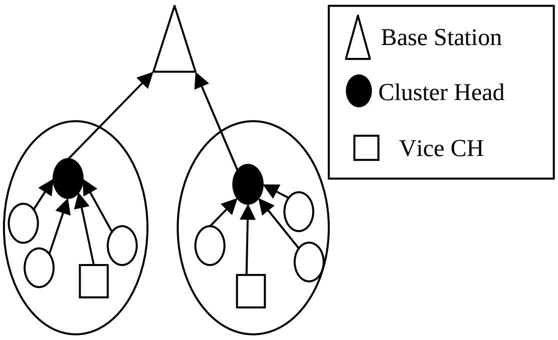


Fig. 4 V-LEACH Protocol

F. Cell-LEACH

In this proposed leach protocol, network is divided in different sections called cell. Each cell contains several sensors and one sensor is selected as cell head. Each seven nearby cells form a cluster, with a sensor known as cluster-head. Clustering will remain same as long as network is working; only cell-heads and cluster-heads change randomly. In each cell, Cell-head assign a time slice based on TDM (Time Division Multiplexing) to sensor nodes. Each cell should send its data to the cell-head in allocated time only. This method would also use for sending data from cell-head to CH. When sending information, all the nodes stay off (except the node that is already assigned slicing time). Then cell head will either remove duplicate data or aggregate received data from different sources. After deleting duplicate data and aggregate information in cell-head, this information will be send to cluster-heads and also all the functions in cell-head will be performed as well. The same technique will be used for the selection of cell-head and cluster-head. Initially after the network setup, a cell-head within each cell and a cluster-head within each cluster will be selected dynamically, because all the sensor nodes have the same energy. In next round, each old cell-head allow to choose a new cell-head dynamically and replace it with new one [5].

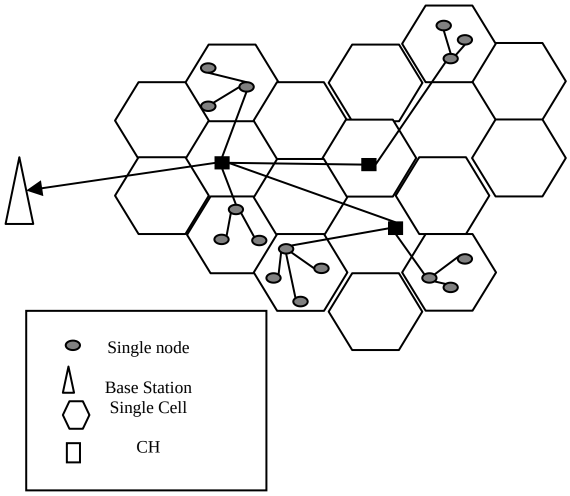


Fig. 5 Cell LEACH

G. *LEACH-F (Fixed No. of Clusters Low Energy Adaptive Clustering Hierarchy)*

In LEACH-F, once the clusters are formed they are fixed and there is no network setup overhead at the beginning of each round. To make the decision for clusters, LEACH-F uses the same centralized cluster formation algorithm as LEACH-C. In LEACH-F, new nodes cannot be added to the system and do not adjust their behaviour when any node dies in the network. Furthermore, the node mobility cannot be handled by the LEACH-F. Only the cluster head position is rotated among the nodes within the cluster same as LEACH protocol [4]. The advantage of this process in comparison to LEACH is that, there is no network set-up overhead at the beginning of each round. LEACH-F may or may not be provided energy saving. A stable cluster and rotating cluster head concept is followed by LEACH-F in which once cluster formed is maintained stable throughout the network lifetime in order to avoid re-clustering [8].

H. *LEACH-L (Energy Balanced Low Energy Adaptive Clustering Hierarchy)*

Leach-L is an advanced multi-hop routing protocol and based only on the distance. It is best suited for large area wireless sensor network and the optimum hop counts are deducted. When the base station is located close to the CH then CH can communicate directly to the BS, but when they are at far the distance from the base station, they can communicate with each other by the multi-hop way and the shortest transmission distance is limited. Sensors are allowed to use different frequencies to communicate with the base station. In each round clusters are updated where each round has two phases: set-up phase and steady phase. In each round new cluster head is chosen and the load is distributed and balanced among all the nodes of the network. Since Leach-L distribute power equally among all sensors, in the pre-period, the network's activity nodes and cover areas of Leach- L is greatly larger than that of Leach-M [4], [8].

I. *LEACH-B (Balanced Low Energy Adaptive Clustering Hierarchy)*

LEACH-B is an enhancement of original Leach Protocol known as Balanced-LEACH uses the decentralized algorithms of cluster formation in which each sensor node has knowledge only about its own position and the final receiver or destination and has no information about the location of all the sensor nodes. Leach-B includes the following techniques:

1. An Algorithm for Cluster head selection
2. Cluster formation
3. And transmission of data with multiple access

By analysing the energy lost in the path between final receiver and itself, each of the sensor nodes chooses its cluster head. Efficiency of Leach-B is better than Leach [4].

J. *LEACH-A (Advanced Low Energy Adaptive Clustering Hierarchy)*

In original Leach protocol, Cluster Head is responsible for transmitting data to BS directly which consumes high energy than other member nodes in the network [3]. Hence both the energy saving is improved and reliable data transfer in LEACH-A. In Advanced-LEACH, the data is processed by using a technique called mobile agent. Advanced Leach may be defined as a heterogeneous energy protocol is developed for the purpose of energy saving, reliable data transfer, decreasing the probability of node's failure and for increasing the time interval before the death of the first node. It uses synchronized clock, through which each sensor gets the starting of each round [4], [8].

Following are the advantages of Leach-A protocol:

1. The collaboration of data reduces the amount of information that is transmitted to the BS.
2. It uses TDMA/CDMA techniques which allow hierarchy and does clustering on different levels which can save more energy.

Functions performed by the gateways:

1. Energy consumption decreases.
2. Lifetime of the cluster head extends.
3. Reduction in the nodes failure probability.
4. Extends the time interval before the death of the first node.
5. Increasing the overall lifespan of WSNs.

K. LEACH-M (Mobile-Low Energy Adaptive Clustering Hierarchy)

Mobility support is a fundamental issue in the Leach routing protocol. Leach-M is designed for this issue. In Leach-Mobile protocol, cluster head nodes and non cluster head nodes can move during the set-up and steady phase [7]. In Leach-M the nodes are homogeneous and obtain their location information through GPS and Base station is considered to be fixed [4]. To select appropriate cluster head Distributed setup phase of LEACH is modified by M-LEACH. The optimum cluster head can be selected on the basis of minimum mobility and lowest attenuation mode, which broadcast their status to all nodes which are in its transmission range. Another criteria for the selection of the cluster-head is mobility speed. In the steady state phase of the original LEACH protocol, another cluster head is chosen if nodes move away from the cluster-head or cluster-head moves away from its member nodes which results into inefficient clustering formation. To tackle this problem M- LEACH provides a handover mechanism for nodes to switch on to new cluster-head [4], [7].

L. LEACH-S (SOLARAWARE CENTRALIZED LEACH)

Energy harvesting is an essential application in some wireless sensor network, especially when nodes are placed in areas which are non-accessible. For such kind of applications, solar-aware LEACH (LEACH-S) has been proposed in which solar power can extend the lifetime of the wireless sensor network. Both LEACH and LEACH-C is the extension of LEACH-S [7].

1) Solar-Aware Centralized LEACH

In LEACH-S, base station uses improved central control algorithm to select the cluster head. Base station normally selects solar powered nodes that have maximum residual energy. In LEACH-S, nodes transmit its solar status to the base station along with the energy and nodes with higher energy are selected as the cluster head. The performance of the sensor network increases with the increase in solar-aware nodes. The sun duration prolongs the lifetime of the sensor network. If the sun duration is smaller cluster head handover is done in LEACH-S [7], [8].

2) Solar-aware Distributed LEACH

In this LEACH-S, choosing preference of cluster head is given to solar driven nodes whose probability is higher than battery-driven nodes [4].

TABLE: COMPARISION OF LEACH AND ITS DESCENDENT PROTOCOLS

LEACH and its Descendant	**Abbreviation**	**Differ from LEACH**
E-LEACH	Energy LEACH	Selection of ch is based on Residual Energy.
TL-LEACH	Two Level LEACH	Ch sends the Data to bs through a Ch that lies between the Ch and bs.
M-LEACH	Multihop LEACH	CH Relays the data to BS through Multiple CH as Relay Nodes.
C-LEACH	Centralized LEACH	BS is Responsible for making Clusters for each round by Running Centralized Cluster Algorithm by Getting Remaining Energy and Position of each Sensor Node.
LEACH-MOBILE	LEACH-MOBILE	LEACH MOBILE is best Suited for Mobility Centric Environment.
V-LEACH	Vice Cluster Head LEACH	There is a vice-CH that play the role of the CH when the CH dies.
CELL-LEACH	Cell LEACH	Sensor Network is Divided in Different Sections Which are called Cell.
LEACH-A	ADVANCED LEACH	LEACH-A Provides Reliable Data Transfer in Network.
LEACH-B	BALANCED LEACH	LEACH-B Choose its CH by Calculating the Energy need for the Path Between itself and Destination.
LEACH-S	Solar Aware LEACH	Selection of CH is based on Residual Energy Level.
LEACH-L	Advanced Multihop LEACH	CH Selection is based on Distance.

IV. CONCLUSION

In this paper, the most important challenge in designing routing protocols for Wireless Sensor Networks is energy efficiency, which is due to the limited energy of the sensors. The most important goal of a routing protocol is to increase the lifetime of sensor nodes. Sensors mainly consume energy during data transmission and reception. Therefore, routing protocols should be energy efficient to enhance not only the individual node lifetime, but also extend the lifetime of the whole of the wireless sensor networks. Therefore, because of this reason LEACH protocol is selected. It gives better performance in both the energy efficiency and the network lifetime. We can say the advantages of LEACH overcome the problem of WSN. Along with the advantages of LEACH it also has some disadvantages. Thus to overcome from these disadvantages and to make it energy efficient many descendants of LEACH protocol are introduced and some of them like E-LEACH, TL-LEACH, M-LEACH, LEACH-C, CELL-LEACH and V-LEACH are described in this paper. This descendant or improved LEACH gives better result than normal LEACH. Each of the descendant routing protocols has its own advantages as compared to the fundamental one.

REFERENCES

[1] Lan Tien Nguyen †1, Xavier Defago †2, Razvan Beuran *†3, Yoichi Shinoda †*4 IEEE ISWCS 2008.

[2] Nikolaos A. Pantazis, Stefanos A. Nikolidakis and Dimitrios D. Vergados, Senior Member, IEEE, IEEE Communications Surveys & Tutorials, Vol. 15, No. 2, Second Quarter 201

[3] M. Usha1, Dr. N. Sankarram2, International Journal of Innovative Research in Computer and Communication Engineering Vol. 2, Special Issue 1, March 2014.

[4] J. Gnanambigai1, Dr. N. Rengarajan2, K. Anbukkarasi3, "Leach and Its Descendant Protocols: A Survey" International Journal of Communication and Computer Technologies Vol. 01, No. 3, Issue: 02 September 2012.

[5] Ravneet Kaur1, Deepika Sharma2, Navdeep Kaur3, "Comparative Analysis Of Leach and Its Descendant Protocols In Wireless Sensor Network" International Journal of P2P Network Trends and Technology-Vol. 3, Issue 1–2013

[6] Sapna Choudhary1, Sandeep Sharma2, International Journal of Advanced Research in Computer Science and Software Engineering Vol. 4, Issue 1, January 2014

[7] M. Aslam1, N. Javaid2, A. Rahim3, U. Nazir4, "Survey of Extended LEACH-Based Clustering Routing Protocols for Wireless Sensor Networks" [cs.NI] 11 July 2012.

[8] P. Manimala1, R. Senthamil selvi2, International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250–2459, ISO 9001:2008 Certified Journal, Vol. 3, Issue 12, December 2013).

Performance Research on Firefly Optimization Algorithm with Mutation

Sankalap Arora[1] and Satvir Singh[2]
[1]*Department of Computer Science & Engineering,*
SUS College of Engineering and Technology, Mohali, India
[2]*Department of Electronics & Communication Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozpur, India
E-mail: [1]sankalap.arora@gmail.com, [2]drsatvir.in@gmail.com

***Abstract*—Firefly algorithm is an optimization algorithm which mimics the behavior of fireflies to solve problems. In this paper, firefly algorithm with mutation is researched and the performance effect of parameter settings is studied in order to show which setting is more suitable for solving optimization problems. It is tested on ten standard function problems and compared with original firefly algorithm. Experiment results show that firefly with mutation is effective for solving most of the benchmark functions. And the firefly algorithm with mutation has superior performance to the compared method on all ten standard benchmark functions.**

Keywords—Optimization, Firefly, Mutation, Algorithm, Performance

I. Introduction

An optimization problem is the problem of finding the best solution from all feasible solutions. Classical methods of optimization are generally not used for their impracticality in complicated real life situation. They are generally deterministic in nature. Nature-inspired metaheuristic algorithms such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Ant Colony Optimization (ACO) and Firefly Algorithm (FA) are most powerful algorithms for optimization [1]. The goal is to develop more proficient and better optimization techniques that might involve more and more sophistication of algorithm. Nature has remained a great source of inspiration to mankind to develop novel methods of optimization techniques. Bio mimicking of several natural events have given birth to modern day metaheuristic algorithm. The main essence of metaheuristic algorithm is to exploit the method of trial and error. Meta-heuristics have been remarkably successful because of four main reasons: simplicity, flexibility, derivation free mechanism, and local optima avoidance [2].

First, meta-heuristics are fairly simple as they are inspired by very simple concepts. The inspirations are typically related to physical phenomena, animals' behaviors, or evolutionary concepts [3]. The simplicity allows researchers to simulate different natural concepts, propose new meta-heuristics, hybridize two or more meta-heuristics, or improve the current meta-heuristics. Moreover, the simplicity assists researchers to learn metaheuristic quickly and apply them to their problems [4], [5].

Second, flexibility refers to the ease of applicability of metaheuristics to different problems without any major changes in the algorithm. Meta-heuristics are readily applicable to different problems since they mostly assume problems as black boxes. In other words, only the input(s) and output(s) of a system are important for a meta-heuristic which change according to the problem.

Third, in contrast to gradient-based optimization approaches, meta-heuristics optimize problems stochastically [6]. The optimization process starts with random solution(s) and there is no need to calculate the derivative of search spaces to find the optimum. This makes meta-heuristics highly suitable for real world problems [7], [8].

Lastly, meta-heuristics have capability to avoid local optima because of the stochastic nature of meta-heuristics which allow them to avoid stagnation in local solutions and search the entire search space extensively. The search space of real problems is usually unknown and very complex with a massive number of local optima, so meta-heuristics are good options for optimizing these challenging real world problems [9].

The strength of standard firefly algorithm lies in the attractiveness of less brighter firefly towards the brighter firefly [10]. The less brighter firefly improvises its position according to brighter firefly but it does not add good features or attributes from the better firefly. So if the less brighter firefly can add features or attributes from the better firefly, it can converge to optima quickly in less number of iterations. [11]

This paper aims to research on Firefly Algorithm with mutation (MFA) and provide comparison study of the MFA with FA. We will first outline the Firefly Algorithm, then formulate the firefly algorithm with mutation and then demonstrate the comparison of these algorithms focusing on critical factors like convergence and time consumption. The MFA optimization seems more promising in the sense that MFA converges quickly than firefly algorithm optimization. [12]

II. Firefly Algorithm

A. *Standard Firefly algorithm*

1) *Behavior of fireflie*

There are around two thousand species of firefly algorithm, usually found in tropical and temperate regions. Most species of fireflies produce unique, short and rhythmic flashes. Bioluminescence is the process responsible for flashing of light. These flashes are used to attract mating partners and potential prey. These rhythmic flashes are different from each other on the basis of the rate of flashing and amount of time. Females respond to unique pattern of flashing of a male which forms a signal system bringing both sexes together. [13].

When a light source emits light intensity from a particular distance r it obeys inverse square law. The light intensity I decreases with increase in the distance r in terms of $I \propto 1/r^2$ [14]. Air acts as an absorbent and light becomes weaker as the distance increase [15]. The flashing light is formulated in such a way that it can associated with objective function to be optimized, which opens gateway in formulation of new optimization algorithms [10].

2) Firefly Algorithm

To develop firefly inspired algorithm, it is mandatory to idealize some of the characteristics of fireflies. For simplicity in describing Firefly algorithm, three assumptions have been made [16]:

1. All fireflies are of same sex which means every firefly will be attracted to other fireflies regardless of their sex.
2. Attractiveness is proportional to their brightness, thus for any two flashing fireflies, the less brighter one will movetowards the brighter one.
3. The brightness of a firefly is affected or determined by the landscape of the objective function. Based on these three rules, the basic steps of the firefly algorithm (FA) can be summarized as the pseudocode shown in Algorithm 1 [17].

Algorithm 1 Firefly algorithm

```
Objective Function f (X), X= $(x_1, x_2 \dots\dots x_d)$
Generate the initial population of n fireflies, $X_i$,
i = 1, 2,…, n
Light intensity $I_i$ at $X_i$ is determined by f $(X_i)$
Define the light absorption coefficient γ
while (t < MaxGeneration)
        for= 1: $n$, all n fireflies
                for j= 1: $n$,all n fireflies (inner loop)
                if $(I_i < I_j)$, Move firefly $i$ towards $j$;
                end if
    Vary attractiveness with distance $i$ via exp[− γr$^2$]
                end for $j$
        end for $i$
    Rank the fireflies and find the current global best
solution g*
    end while
    Post-process the results
```

3) Attractiveness and Distance

Firefly algorithm is based on two important factors: variation of light intensity and formulation of the attractiveness. An assumption is made that attractiveness of a firefly is calculated according its brightness which is associated further with the encoded objective function [18]. In maximum optimization problems, the brightness of a firefly can be chosen as $I(x) \propto f(x)$where I is the intensity of a firefly and x is a particular location. Attractiveness β is relative and it will change according to distance r_{ij} between firefly i^{th} and fireflyj^{th}. Light is also absorbed by the air and it also get decreased with increasing distance so attractiveness is allowed to vary with degree of absorption. Light intensity $I(r)$ varies according to inverse square law and for a given medium with fixed light absorbtion coefficientγ the light intensity I varies with distance r [19].So attractiveness β of firefly is defined by

$$\beta = \beta_o \exp(-\gamma r^2) \qquad (1)$$

where β_o is the attractiveness at distance $r = 0, \beta$ is the fixed light absorption coefficient for a specific medium and γ is light absorption coefficient. The distance r_{ij} between any two fireflies i^{th} and j^{th} located at X_i and X_j, respectively, is determined using the Euclidean norm and movement of a less brighter fireflyi^{th} towards brighter firefly j^{th} is determinedby

$$x_i = x_i + \beta_0^{e\left(-\gamma r_{ij}^2\right)}\left(x_j - x_i\right) + \alpha\left(rand - \frac{1}{2}\right) \qquad (2)$$

In (2) the second term is due to relative attraction and third term is a randomization parameter. α is randomization parameter normally selected within range [0,1] and $rand$ is a random number uniformly distributed in [0, 1]. Now to introduce the variation of attractiveness, γ parameter is used and its range is 0.01 to 10. The initial locations of n fireflies are distributed uniformly in the search space whenever the number of fireflies are greater than number of local optima [20]. During the execution, the fireflies converge into all of these local optima, the global optima is determined. The algorithm will approach the global optima when $n \to \infty$ and number of iterations are greater than 1 but in reality it converge quickly [21].

In (2) the second term is due to relative attraction and third term is a randomization parameter. $\alpha$is randomization parameter normally selected within range[0,1] and ϵ_i is a vector of random numbers drawn either a Gaussian or uniform distribution. Now to introduce the variation of attractiveness, γ parameter is used and its range is 0.1 to 10. In optimization problem where number of fireflies are greater than number of local optima, the initial locations of the *n* fireflies should be distributed relatively uniformly throughout the entire search space. During the execution, the fireflies converge into all of these local optima, the global optima is determined. FA will approach the global optima when $n \to \infty$ and number of iterations are greater than 1 but in reality it converge extremely quickly.

B. Firefly Algorithm with Mutation

The strength of any optimization algorithm lies in how faster the algorithm explores the new possible solutions and how efficiently it exploit the solutions to make them better. FA algorithm performs a move step which contains the exploration and exploitation concept. There is a need by which exploration and exploitation can be enhanced and the algorithm can work more efficiently. So mutation is added to firefly algorithm to achieve better results. By using mutation the basic concept of searching solutions is modified. In standard firefly algorithm, space is searched by moving the less brighter firefly moves towards the more brighter firefly. Firefly algorithm with mutation searches the search space by adding features to less brighter firefly from more brighter firefly. The extent of features to be added is decided by calculating the mutation probability of each firefly. The better the firefly, lesser the mutation probability and viceversa. By using features of better fireflies, the algorithm will converge faster and avoid falling into the local optimum.

In the firefly algorithm with mutation all fireflies do not participate in mutation, but some. The underlying principle of MFA is to adapt features from other fireflies and achieve best values in minimum amount of time. The mutation concept is also modified for FA. To better explain it lets suppose there are 100 fireflies, only top good 40 percent individuals will donate their features because they are good because of their good features. Similarly, the need of the good features from top 40 per cent will be needed by last 40 per cent of fireflies i.e. worst 40 percent of solutions. The better the firefly, more the mutation probability and worse the firefly, less the mutation probability. Additional to it, the in between 20 percent individuals which are average i.e. neither good nor bad, do not participate in the mutation process and they have very low mutation probability. The basic principle which is followed in the firefly algorithm with mutation is that there is better probability of good solutions becoming better and there is low probability of bad solutions becoming very good. So to make bad solutions, better solutions, mutation can be applied. As mutation gave them chance to modify themselves and attain good features.

The mutation operator is used to change some elements in selected individuals with a probability p_m (mutation probability) leading to additional firefly diversity to help the search process escape from local optimal traps. Each firefly has its mutation probability$p_m$necessary for it to mutation. The choice of p_m will critically affect the behavior and performance. Typical values of p_m are same as in GA i.e. 0.001 to 0.05. The mutation probability (MP) in firefly algorithm with mutation is calculated using (3)

$$MP = f_{new} - f_{old} \tag{3}$$

where f_{new} is the fitness of the new firefly and f_{new} the fitness of the original firefly. For a generation that undergoes n_m mutation operations, the average mutation progress value $\acute{M}P$ is given by (4)

$$\acute{M}P = \frac{1}{n_m}\sum MP \tag{4}$$

Before the end of each generation, mutation rates are adjusted using these average progress values. Based on these equations (3) and (4), the steps of the firefly algorithm with mutation (MFA) can be summarized as the pseudocode shown in Algorithm 2.

Algorithm 1 Firefly algorithm with Mutation

```
Objective Function f (X), X= $(x_1, x_2 \ldots \ldots . x_d)$
Generate the initial population of n fireflies, $X_i$,
i = 1, 2,…, n
Light intensity $I_i$ at $X_i$ is determined by f $(X_i)$
Define the light absorption coefficient γ
while (t < MaxGeneration)
        for= $1{:}\,n$, all n fireflies
        for j= $1{:}\,n$,all n fireflies (inner loop)
        if $(I_i < I_j)$, Move firefly $i$ towards $j$;
        Calculate Mutation probability
        Perform Mutation
        end if
        Vary attractiveness with distance $i$ via exp[− γr$^2$]
        end for $j$
        end for $i$
    Rank the fireflies and find the current global best solution g*
    end while
    Post-process the results
```

III. Simulation & Experiments

In this paper, ten standard benchmark functions are used for testing the success of firefly algorithm with mutation against standard firefly algorithm, which are described in Table I.

There are many ways to carry out the comparison of algorithm performance and two the obvious approaches are: to compare the numbers of function evaluations for a given tolerance or accuracy or to compare their accuracies for a fixed number of function evaluations. Here the second approach is used. In simulations, maximum number of evaluations is fixed to 1000, so that meaningful statistical analysis can be done.

For both the algorithms, same standard of learning parameters i.e. $\alpha = 0.1$ and $\gamma = 0.01$ are used. For better comparison between the two algorithms Mean, Median, Best values and Worst values for different n are also being considered. Different number of fireflies(n) are used having values 50, 100, 250 and 500. Number of dimensions was as per the standard benchmark function, as discussed in Table I.

TABLE I VARIOUS STANDARD BENCHMARK FUNCTIONS

Benchmark Function	Formula	Dimension (n)	Range	Optimal Value
Ackley	$f(\vec{x}) = -20.exp\left({}^{-0.2}\sqrt{\frac{i}{n}\sum_{i=1}^{n} x_i^2}\right) - exp\left(\frac{1}{n}\sum_{i=1}^{n} cos(2\pi.x_i)\right) + 20 + e$	30	(-32,32)	0
Sphere	$f_0(\vec{x}) = \sum_{i=1}^{n} x_i^2$	30	(-100,100)	0
Griewank	$f(\vec{x}) = \frac{1}{4000}\sum_{i=1}^{n-1}(x_i - 100)^2 - \prod_{i=1}^{n-1}\cos\left(\frac{x_i - 100}{\sqrt{i-1}}\right) + 1$	10	(-600, +600)	0
Michalewiz	$f(x) = -\sum_{i=0}^{n}(\sin(x_i)sin^{20}\left(\frac{ix_i^2}{\pi}\right)$	10	(0,π)	-0.966n
Rastrigin	$f(\vec{x}) = \sum_{i=1}^{n}(x_i^2 - 10\cos(2\pi x_i) + 10)$	30	(-5.12, 5.12)	0
Schaffer	$f(x) = (x_0^2 + x_1^2)^{\frac{1}{4}}(50(x_0^2 + x_1^2)^{0.1} + 1)$	2	(-100,100)	0
Schewel	$\sum_{i=0}^{n-1}\lvert x_i\rvert + \prod_{i-0}^{n-1}\lvert x_i\rvert$	30	(-10,10)	0

These algorithms are implemented in QT Creator. Table II shows best solution, mean solution, median solution, worst solution and time taken to complete defined number of iterations. Table II reveals the striking potential of the MFA in obtaining the high precision optimal solutions better than the standard FA solutions.

TABLE II VARIOUS STANDARD BENCHMARK FUNCTIONS

Functions	No. of Particles	Comparison of Simulation Results								Time Taken	
		Mean		Median		Best Results		Worst Results			
		FA	MFA	FA	MFA	FA	MFA	FA	MFA	FA	MFA
Ackley	50	9.89E-04	9.11E-04	9.85E-04	8.83E-04	9.01E-04	7.93E-04	1.13E-03	1.05E-03	4	1.3
	100	8.86E-04	7.66E-04	8.80E-04	7.60E-04	8.08E-04	6.75E-04	9.66E-04	8.27E-04	9	4.7
	250	8.19E-04	7.05E-04	8.54E-04	7.16E-04	7.47E-04	5.75E-04	9.03E-04	7.61E-04	37	25.3
	500	7.42E-04	6.63E-04	7.26E-04	6.48E-04	6.97E-04	6.28E-04	8.07E-04	6.99E-04	125	98
Sphere	50	1.91E-05	1.45E-05	1.90E-05	1.40E-05	1.60E-05	1.20E-05	2.20E-05	1.70E-05	1	1
	100	1.43E-05	1.21E-05	1.40E-05	1.20E-05	1.10E-05	1.00E-05	1.70E-05	1.60E-05	4	3.9
	250	1.19E-05	9.30E-06	1.20E-05	1.00E-05	1.00E-05	7.00E-05	1.40E-05	1.10E-05	29	24
	500	9.60E-06	8.90E-05	1.00E-05	9.00E-05	7.00E-06	7.00E-06	1.10E-05	1.10E-05	120	98
Griewank	50	1.44E-01	1.24E-01	9.35E-02	1.79E-01	3.60E-02	2.04E-01	4.06E-01	1.03E-01	1	0.55
	100	1.49E-01	1.51E-01	1.03E-01	7.63E-02	4.19E-02	2.36E-01	4.36E-01	1.75E-01	3	2.03
	250	1.53E-01	1.54E-01	1.40E-01	6.40E-02	4.67E-02	1.99E-01	3.91E-01	1.48E-01	12	9.6
	500	1.74E-01	1.44E-01	1.50E-01	2.12E-01	8.37E-02	1.94E-01	3.62E-01	1.23E-01	58	44.7
Michalewiz	50	-7.34E+00	-8.88724	-7.31E+00	-9.2088	-8.65E+00	-9.49491	-4.28E+00	-8.16183	2	0.62
	100	-7.86E+00	-8.89E+00	-8.13E+00	-9.21E+00	-8.70E+00	-9.49E+00	-6.16E+00	-8.16E+00	5	2.04
	250	-7.78E+00	-8.30E+00	-7.94E+00	-8.45E+00	-9.03E+00	-9.29E+00	-6.67E+00	-6.63E+00	15	11.8
	500	-7.56E+00	-8.88E+00	-7.64E+00	-9.40E+00	-8.40E+00	-9.02E+00	-7.11E+00	-9.26E+00	58	45.8
Rastrigin	50	4.71E+01	4.26E+01	4.58E+01	4.38E+01	2.79E+01	2.48E+01	6.17E+01	6.07E+01	2	1.2
	100	4.97E+01	4.38E+01	3.98E+01	3.98E+01	2.98E+01	2.49E+01	8.76E+01	6.57E+01	7	4.3
	250	4.32E+01	4.09E+01	4.08E+01	3.28E+01	3.18E+01	2.49E+01	6.57E+01	7.96E+01	32	25.4

Table II (Contd.)…

…Table II (Various Standard Benchmark Functions)

Functions	**No. of Particles**	**Comparison of Simulation Results**								**Time Taken**	
		Mean		**Median**		**Best Results**		**Worst Results**			
		FA	**MFA**	**FA**	**MFA**	**FA**	**MFA**	**FA**	**MFA**	**FA**	**MFA**
	500	4.01E+01	4.05E+01	3.68E+01	3.38E+01	2.89E+01	2.29E+01	5.27E+01	6.37E+01	122	99
Schaffer	50	2.76E-02	3.03E-02	3.32E-02	2.93E-02	1.76E-02	2.68E-02	4.00E-02	3.23E-02	0.7	0.2
	100	1.90E-02	2.60E-02	1.74E-02	2.53E-02	6.50E-03	1.95E-02	3.32E-02	3.90E-02	1.5	0.8
	250	1.61E-02	1.59E-02	1.54E-02	1.38E-02	4.66E-03	8.43E-03	2.47E-02	2.84E-02	7.5	5.1
	500	1.65E-02	1.68E-02	1.52E-02	1.36E-02	8.51E-03	7.09E-03	2.53E-02	2.92E-02	22	21.1
Schewel	50	1.91E-03	1.73E-03	1.85E-03	1.70E-03	1.79E-03	1.44E-03	2.09E-03	2.00E-03	4	1
	100	1.79E-03	1.50E-03	1.76E-03	1.51E-03	1.41E-03	1.36E-03	2.13E-03	1.61E-03	9	3.9
	250	1.53E-03	1.29E-03	1.54E-03	1.34E-03	1.36E-03	1.02E-03	1.74E-03	1.47E-03	36	24.4
	500	1.44E-03	1.28E-03	1.41E-03	1.27E-03	1.34E-03	1.19E-03	1.51E-03	1.36E-03	129	103

IV. Conclusion and Future Scope

In the paper, firefly algorithm with mutation(MFA) is researched which considers mutation probability and then perform mutation on fireflies to better explore search space. Simulation results demonstrated the potential of MFA. Simulation results suggests that the proposed algorithm is superior to standard firefly algorithm in terms of both efficiency and success rate. The standard firefly algorithm is efficient but solutions still change as the optima are approaching. So the solution quality is improved by reducing randomness by introducing mutation probability concept. Further, convergence is also improved. The reason for these better results lies in the modified mutation concept which gave bad fireflies/solutions more chance to adapt from good fireflies and eventually become better in less amount of time. The amount of time consumption is less because of the fact that all solutions do not participate in mutation but only those who can donate good features and those who need good features which is calculated by mutation probability for each individual. Considering more iterations information of the algorithm and its application in combination with other algorithms could be an exciting direction in the future.

References

[1] E. Bonabeau, M. Dorigo, and G. Theraulaz, Swarm intelligence. Oxford, 1999.

[2] D.E. Goldberg and J.H. Holland, “Genetic algorithms and machinelearning,” Machine learning, Vol. 3, No. 2, pp. 95–99, 1988.

[3] V. Gazi and K.M. Passino, “Stability analysis of social foragingswarms,” Systems, Man, and Cybernetics, Part B: Cybernetics, IEEETransactions on, Vol. 34, No. 1, pp. 539–557, 2004.

[4] J. Kennedy and R. Eberhart, “Particle swarm optimization,” 1995.

[5] K.M. Passino, “Biomimicry of bacterial foraging for distributed optimizationand control,” Control Systems, IEEE, Vol. 22, No. 3, pp. 52–67, 2002.

[6] T. B¨ack, D.B. Fogel, and Z. Michalewicz, Evolutionary computation 1: Basic algorithms and operators. CRC Press, 2000, Vol. 1.

[7] R.C. Eberhart, Y. Shi, and J. Kennedy, Swarm intelligence. Elsevier, 2001.

[8] J. Liang, P. Suganthan, and K. Deb, “Novel composition test functionsfor numerical global optimization,” in Swarm Intelligence Symposium, 2005. SIS 2005. Proceedings 2005 IEEE. IEEE, 2005, pp. 68–75.

[9] X.S. Yang, Nature-inspired metaheuristic algorithms. Luniver press, 2010.

[10] S. Łukasik and S. Zak, “Firefly algorithm for continuous constrainedoptimization tasks,” in Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems. Springer, 2009, pp. 97–106.

[11] X.S. Yang, Engineering optimization: an introduction with metaheuristicapplications. John Wiley & Sons, 2010.

[12] S. Arora and S. Singh, “The firefly optimization algorithm: Convergenceanalysis and parameter selection,” International Journal of ComputerApplications, Vol. 69, No. 3, pp. 48–52, 2013.

[13] X.S. Yang, “Firefly algorithm, levy flights and global optimization,”in Research and Development in Intelligent Systems XXVI. Springer, 2010, pp. 209–218.

[14] S. Arora and S. Singh, “A conceptual comparison of firefly algorithm, batalgorithm and cuckoo search,” in Control Computing Communication & Materials (ICCCCM), 2013 International Conference on. IEEE, 2013, pp. 1–4.

[15] X.S. Yang, “Chaos-enhanced firefly algorithm with automatic parametertuning,” International Journal of Swarm Intelligence Research (IJSIR), Vol. 2, No. 4, pp. 1–11, 2011.

[16] K. Deb, Optimization for engineering design: Algorithms and examples. PHI Learning Pvt. Ltd., 2012.

[17] X.S. Yang, “Firefly algorithm, stochastic test functions and designoptimisation,” International Journal of Bio-Inspired Computation, Vol. 2, No. 2, pp. 78–84, 2010.

[18] T. Apostolopoulos and A. Vlachos, “Application of the firefly algorithmfor solving the economic emissions load dispatch problem,” International Journal of Combinatorics, Vol. 2011, 2010.

[19] X.S. Yang, “Firefly algorithms for multimodal optimization,” in Stochastic algorithms: foundations and applications. Springer, 2009, pp. 169–178.

[20] Z.W. Geem, J.H. Kim, and G. Loganathan, “A new heuristic optimizationalgorithm: harmony search,” Simulation, Vol. 76, No. 2, pp. 60–68, 2001.

[21] X.S. Yang, “Biology-derived algorithms in engineering optimization,” arXiv preprint arXiv: 1003. 1888, 2010.

A Competent Study of Hybrid Routing Protocols of MANETs Using NS-2 Simulator

Rohit Kumar[1], Meenakshi Sharma[2], Navdeep Kaur[3] and Gurjeevan Singh[4]
[1,2,3,4]Department of Electronics & Comm. Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab
E-mail: [1]rk74447@gmail.com, [2]meena.leo10@gmail.com, [3]navdeepkaurjhaj213@gmail.com,
[4]gurjeevansandhu@gmail.com

***Abstract*—An ad-hoc network refers to any set of networks where all devices have equal status on the network and are free to associate with any other adhoc network devices in link range. Mobile Adhoc Networks (MANETs) are an integral part of next generation networks because of its flexibility, ease of maintenance, infrastructure less nature, self administration capabilities, auto configuration and cost effectiveness. Various researchers in MANETs focus on proactive and reactive routing protocols. But this paper focuses on the combination of these two i.e. the third type of routing protocol i.e. hybrid protocols. The three types of hybrid protocols considered in this paper are TORA, ZHLS and ZRP. The performance metrics used for comparison purpose are routing overhead, network overload and average end to end delay.**

Keywords: MANET, TORA, ZHLS, ZRP

I. INTRODUCTION

An ad hoc network is usually thought of as a network with nodes that are relatively mobile compared to a wired networks. Hence, the topology of the network is dynamic and the changes often unpredictable oppose to the Internet which is a wired network [5]. In Mobile Ad Hoc Networking, the communication does not rely on any existing infrastructure such as dedicated routers, transceiver base stations etc. Mobile devices (e.g. notebook computers, PDAs, cell phones, etc.) with wireless radio equipment are supposed to communicate with each other, without the help of any other (fixed) devices. In order to make it possible, typically each node needs to act as a router to relay packets to nodes out of direct communication area. Under these conditions, routing is much more complex than in conventional (static) networks. Many of the possible solutions are dogged by the characteristics of the media, the conduct of nodes and the data flow [4]. Mobile Adhoc networks are very attractive for tactical communication in military and law enforcement. They are also expected to play an important role in civilian forums such as convention centers, electronic-conferences, and electronic classrooms. Nodes in this network model share the same random access wireless channel [3].

II. ROUTING IN MANETS

The growing interest in Mobile adhoc Network techniques has resulted in many routing protocol proposals [3]. The routing protocols used in MANETs are dissimilar from routing protocols of conventional wired networks. Some of the reasons are scheduled below:

1. Mobility.
2. Limited transmission range.
3. Frequent Route updates.

The performance criterion of nodes in MANETs are diverse than that of wired networks. A few of the performance metrics of MANET routing protocols are shown below:

1. Energy consumption.
2. Route Stability despite mobility.

Routing protocols used for Mobile Adhoc Networks are basically of three types:

1. Proactive Routing Protocols (Table-Driven)
2. Reactive Routing Protocols (On-Demand)
3. Hybrid Routing protocols

This research paper mainly concentrates on hybrid routing protocols which is combination of both the reactive and proactive routing protocols [10].

III. BRIEF OVERVIEW OF ZRP, ZHLS AND TORA

A. *Zone Routing Protocol (ZRP)*

In case of ZRP, a node proactively preserves routes to the destinations within a local neighborhood area which is known as a routing zone. In ZRP, each node retains its zone radius and there is an overlap of neighboring zones. The ZRP maintains routing zones through a proactive component called intra-zone routing protocol (IARP) which is implemented as a modified distance vector scheme. In contrast, the inter-zone routing protocol (IERP) is responsible for attaining routes to destinations which are located outside the routing zone. The IERP employs a query-response mechanism to find out routes on demand [1].

B. *Zone-based Hierarchical Link State (ZHLS) Routing Protocol*

The Zone-based Hierarchical Link State routing (ZHLS) is a type of hybrid routing protocols. In ZHLS, mobile nodes are aware of their physical locations with support from a locating system like GPS i.e. Global Positioning System. Here, the network for ZHLS is divided into non-overlapping zones which are based on the geographical information. ZHLS employs a hierarchical addressing scheme which contains zone ID and node ID. A node verifies its zone ID according to its location and the pre-defined zone map is renowned to all nodes within the network. It is understood that a virtual link connects two zones if at least a single physical link between the zones is present. A two-level network topology configuration is defined in ZHLS, the node topology and the zone topology [2].

C. Temporary Ordered Routing Algorithm (TORA)

TORA is a kind hybrid protocol, which is dispersed and routers only preserve information about neighboring routers. TORA has the unique property of being highly adaptive and quick in route repair during link failure and it provides multiple routes to destination node. It does not always execute a shortest path calculation and the metric used to launch the routing structure does not signify a distance. It consists of link reversal of the Directed Acyclic Graph (ACG). It makes use of Internet MANET Encapsulation Protocol (IMEP) for link status and Neighbor Connectivity Sensing (NCS). IMEP offers reliable and in-order delivery of all the routing control messages from a node to all of its neighbors, and a notification to the routing protocols each time a link neighbors is formed or busted [7].

IV. Simulation Set Up

The comparative analysis of the routing protocols is performed using NS-2 simulator on Windows-7 operating system. The table listed below describes the hybrid routing protocols used and the conditions specified for simulation purpose.

TABLE I Simulation Arrangement

Routing Protocols Used	TORA, ZHLS and ZRP
Packet Rate	50 packets/ sec
Simulation Area	5Km x5Km
Number of Nodes	25,50,75,100,125,150
Traffic Type	High Quality GSM Voice
Simulation Time	1000 sec.
Node Speed	10m/s
Physical Standard	802.11b

V. Performance Metrics Used

A. Routing Overhead

Ad-hoc networks are intended to be scalable. As the network develops, various routing protocols executes in a different way. The measure of routing traffic raises as the network develops. An important measure of the scalability of the protocol, and the network, is known as routing overhead. It is also defined as the entire number of routing packets transmitted over the network, and is expressed in bits per second (bps) or packets per second (pps) [6].

B. Network Overload

In wireless mobile adhoc networks, when there is congestion in the network due outsized number of nodes which are sending and receiving data beyond the limit of its communication area, this is known as network overload.

C. Average End to End Delay

Average End to End delay of a data packet is time taken by the packets from source node to destination node. Average End to end delay time includes average of all the delays taken by router to seek the path in network consumption, processing delay, propagation delay, and End to end delay for a particular packet which was sent by a pre-specified node, as a source node and received successfully at the destination node is

Average End to end delay, $\mathbf{t_d = t_s - t_e}$

Where t_s is the time when sending of the packet at the pre-specified node starts, and time t_e, is the time when the packet is send by the pre-specified node is received successfully at destination node [9].

VI. Results and Observations

In this research paper, three distinct types of hybrid routing protocols are used for performance evaluation by varying the number of nodes and by keeping the simulation area constant. The performance metrics which are used for discussion purposes i.e. routing overhead, network overload and average end-to-end delay are displayed below graphically.

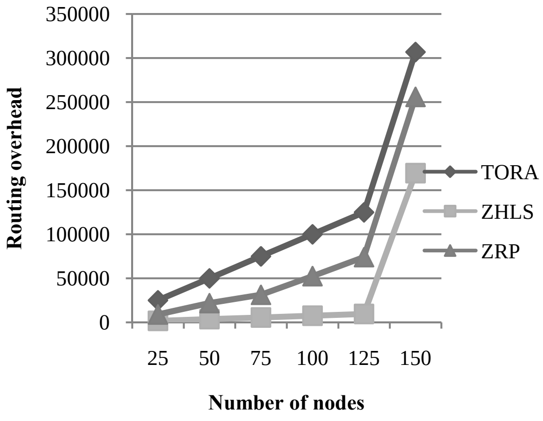


Fig. 1 Routing Overhead for ZRP, ZHLS and TORA by Varying Number of Nodes

Figure 1 illustrates that Routing Overhead for ZHLS is less as compared to ZRP and TORA. Because the number of control packets required by ZHLS is very less in comparison to ZRP and TORA due to the presence of non-overlapping zones. Due to this reason the comparative analysis is in the favor of ZHLS.

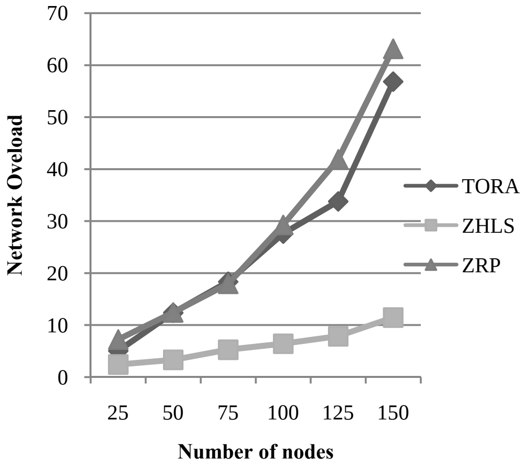


Fig. 2 Network Overload for ZRP, ZHLS and TORA by Varying Number of Nodes

Figure 2 explains that the network overload is less in case of ZHLS than TORA and ZRP i.e. very less congestion is present in case of ZHLS. So, for this performance parameter again the results favor ZHLS.

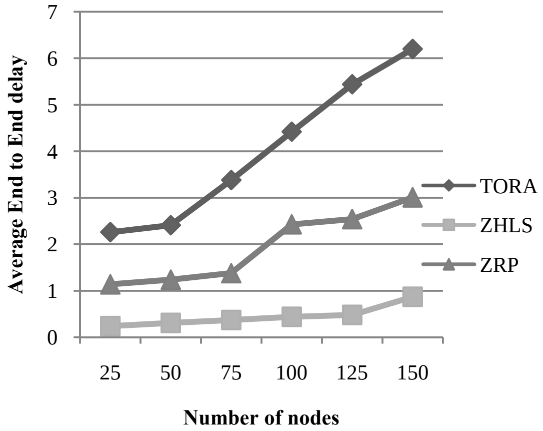


Fig. 3 Average End to End Delay for ZRP, ZHLS and TORA by Varying Number of Nodes

Figure 3 shows that the value of average end to end delay for ZHLS is less in comparison to ZRP and TORA which is one of the important requirements of a routing protocol. So, for a third time for this parameter, ZHLS performs well.

VII. Conclusion

For an adhoc network with large number of nodes which move with different node speed and have different traffic patterns, the hybrid routing protocol is the best selection. In this research paper, when we take review of above discussed performance parameters, ZHLS provides outstanding results than ZRP which is further better than TORA because of least values of routing overhead, network overload and average end to end delay in case of ZHLS.

VIII. Future Scope

In future, this work can be extended by increasing the number of nodes and by increasing the simulation area. Also, the work can be altered by using the other simulators like MATLAB, Glomosim etc.

References

[1] Agrawal, C.d. "Mobile Ad hoc Networking." Centre for Distributed and Mobile Computing, ECECS,University of Cincinnati, 2002.

[2] Changling Liu, J.K."A Survey of Mobile Ad Hoc network Routing Protocols." Universitat Ulm, Fakultat fur Informatik. 2003.

[3] G.V. Sai Aravind, D.S."A Study on Scalable Routing Protocol for Ad Hoc Networks." IJCSET, 2013, pp. 312–317.

[4] Lang, D."A comprehensive overview about selected Ad Hoc Networking Routing Protocols."Department of Computer Science, Technische Universitaat Munchen, Germany, 2003.

[5] Lundberg, D. (2004). "Ad hoc Protocol Evaluation and Experiences of Real World Ad Hoc Networking." Uppsala University, Department of Information Technology, 2004.

[6] Meenakshi Bansal, R.R. "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations." The Internet Society, 1999.

[7] Nadia Qasim, F.S. "Mobile Ad Hoc Networks Simulations Using Routing Protocols for Performance Comparisons." Proceedings of the World Congress on Engineering 2008. London, U.K., 2008.

[8] Navid Nikaein, C.B. "Harp-Hybrid Ad hoc routing protocol. "Proceedings of international symposium on telecommunications, 2001.

[9] Reena, R.P. "Performance Evaluation of Routing Protocols for Manet using NS2." International Journal of Computer Applications, 2013, pp. 12–16.

[10] Seth, A., Seminar Report: Security Issues in MANETs. 2004.

A Survey on Face Detection Techniques

Vikram Mutneja[1] and Satvir Singh[2]
[1,2]Department of Electronics & Communication Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur
E-mail: [1]Vikram.mutneja@gmail.com, [2]DrSatvir.in@gmail.com

***Abstract*—A lot of work has been done till date in the area of face detection. Work on detection of faces has taken leap in various directions since its inception early in the area of image processing and computer vision. The earlier work intended to be focused on detecting faces from still images. With the advent of videos, the focus shifted to detecting and extracting faces from videos. In recent time, with the advent in the surveillance video systems i.e. smart CCTVs, a lot of work has been done in extracting faces from surveillance videos. The main challenges in this area are accuracy and speed. Because of low resolution and small size of faces from surveillance video systems, the accuracy is greatly affected. Further, also the processing of video requires working on high fps (frames per second) videos which affects the speed of operation. So in this paper, we have surveyed the recent work done in the area of detecting and extracting the faces from still images as well as videos, particularly surveillance videos.**

Keywords: Face Detection, Surveillance Video, 3-D Features, CENTRIST

I. INTRODUCTION

Face detection is not a straightforward problem because of factors such as pose variation, occlusions, image orientation, illumination conditions and facial expressions. Face detection techniques can be classified mainly into four categories: Knowledge based methods, Feature-Invariant approaches, Appearance based methods and Template matching methods [1]. Knowledge based methods use the face knowledge to encode rules such as a face often appears in an image with two eyes that are symmetric to each other, a nose and a mouth. Challenge is that it is difficult to translate human knowledge into well defined rules. Feature-invariant approaches extract features such as edges, facial features such as eyes, nose, mouth, and hair line and a statistical model is built to describe their relationships. Challenges in this area are such as lighting, shadows etc. Appearance based methods extract features based upon appearance such as Eigen faces-PCA, Neural Networks, SVM, Ada Boost. In template matching based methods predefined templates have to be stored and correlation values with the standard patterns are computed e.g.: for the face contour, eyes, nose and mouth independently. Limitations so far is that it cannot effectively deal with variation in scale, pose and shape. Challenges are how to represent the template, how to model deformations, efficient matching algorithms etc..

In the surveillance video systems where task is to detect, track and recognize people as well as analyze people activities, detection and extraction of human faces is of paramount importance. It is very important to attach the identity to persons being detected and tracked in the video. From the fact that human faces are used as biometric entity, human faces are generally used to attach identity to a detected human in the surveillance video. Detecting the human faces in surveillance videos is a challenging task on account of various factors such as illumination, low resolution of surveillance cameras, pose variation, facial expressions, face occlusions. The facial images detected from surveillance videos are of very low resolution and not suitable to be applied to face recognition system. Therefore there is also the need to estimate and enhance the quality of detected facial images to bridge the gap between face detection and face recognition systems.

To target the challenges in the area of detecting and extracting facial images from videos, many approaches have been suggested such as incorporating the motion and skin color cues [2] [3], mainly to reduce the search area. Video based techniques work on motion estimation, video object segmentation, background subtraction, skin color detection, object tracking based methods. A recent trend in face detection is to combine multiple information sources such as color, motion, contour etc. More the number of information channels, more will be the accuracy of the system but at the cost of increased detection time.

Inter frame difference technique is most simple and efficient, given (Eq. 1).

$$\Delta(n)_{(x,y)} = I(n)_{(x,y)} - I(n-1)_{(x,y)} \qquad (1)$$

where n represents time and (x, y) the pixel location. The current frame difference $\Delta(n)_{(x,y)}$ is compared with a threshold value, in case it is more than threshold value, the corresponding pixel location is set to 1 otherwise 0. In case of color image, the calculation has to be performed in each color space separately and then aggregated to create final binary image showing the moving parts. Challenges in this area are such as noise removal, threshold estimation etc.

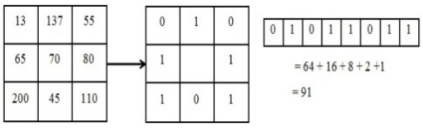
Another category of techniques in motion estimation are optical flow based methods. The biggest advantage of optical flow based techniques is that they can be used even if camera is moving or the background is changing fast. In case of cluttered environment, where background is complex and changing fast, background subtraction technique cannot be used. However optical flow technique has shortcomings as it is more complex, time consuming and poor anti-noise performance.

Skin color cue has also been used in color images to reduce the search area. Difference efficiency has been achieved with various color formats for skin color extraction. RGB format has been found to be best for human vision but not for skin color detection. YCbCr format has been found to give best performance for the skin color detection, as it concentrates the skin colored pixels in a small intensity range. Another challenge in this area is that different cameras produce different colors of same object in scene. So there is also the need of color correction techniques to color of images from different sources.

II. Literature Survey

A lot of work in recent years has been done to develop and use hybrid features in boosting based face detection algorithms. Jun, Bongjin *et al.,* [4] proposed two novel local transform features: Local Gradient Patterns (LGP), modified version of LBP (local Binary Patterns) and Binary Histogram of Oriented Gradients (BHOG), modified version of HOG (Histogram of Oriented Graphics), which were proved to be Faster in computation as compared to LBP and HOG. They proposed hybrid feature that combines various local transform features including LBP, LGP and BHOG by means of AdaBoost method for face and human detection, Hybridization results in improvement of detection performance on account of LBP's robustness to global illumination variations, LGP's robustness to local intensity changes and BHOG's to local pose changes. The proposed local transform features and its hybrid feature have been found to be effective for face and human detection in terms of performance and operating speed.

Some work in recent years has been done to use holistic representation of features for detection, enhancement and recognition of faces. The idea is that the main aim of extracting faces from surveillance videos is to identify the person by applying face recognition. So system efficiency is expected to improve if holistic features are used instead of using separate features for each task such as detection, enhancement and recognition of faces. Bharadwaj, S. *et al.,* [5] have Studied the possibility of using Holistic descriptors Gist and HOG to use in biometric quality assessment of facial images. The spatial properties are preserved in representation, called as Gist. The promising results were obtained in use of above features for quality assessment in face biometrics.



13	137	55
65	70	80
200	45	110

0	1	0
1		1
1	0	1

0	1	0	1	1	0	1	1

=64 + 16 + 8 + 2 +1

=91

Fig. 1 Calculation of 2-D LBP Features

Another dimension in which the work of extracting faces from videos has been addressed is the use of volumetric features i.e., 3-D version of 2-D features such as HAAR, LBP (see Fig. 1), HOG features. The recent work by Martinez-Diaz, Yoanna, *et al.,*. [6] has been done in the use of spatio-temporal based features EVLBP (Extended Volumetric Local Binary Patterns) (see Fig. 2) for detecting the faces from videos. The motivation behind this work was enhancement of efficiency of the system by using 3-D features, which can encode N number of frames in its generation. The main challenge in this area is to select the number of frames, which have to be processed at a time to encode the features. More the value of N, more will be the speed of processing, but accuracy may reduce as it may encode the non facial information while feature encoding. The selection of value of N depends upon how fast the contents of scene in video are changing or how fast the object is moving. Use of EVLBP has been found to give better performance than spatial LBP.

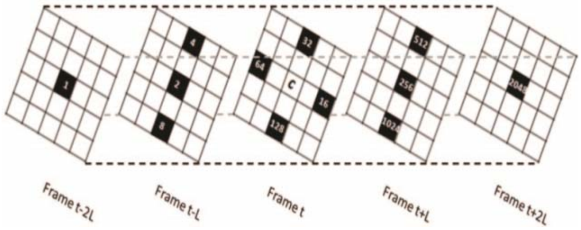


Fig. 2 3-D LBP Features (EVLBP)

Another dimension in which the work of face detection has leapt is the use of parallel and GPU computing [7] [8] [9] [10] [11] [12] to enhance the speed of operation at various stages. The GPU computing has been explored in accelerating the training of boosting based face classifiers. Oro, David, *et al.* [11] presented techniques to increase the performance of the cascade evaluation kernel, which is the most resource-intensive part of the face detection pipeline. Worked on handling problem of GPU underutilization, and achieved a 5X speedup in 1080p videos on average over the fastest known implementations, while slightly improving the accuracy. Also studied the parallelization of the cascade training process and its scalability under SMP platforms. The proposed parallelization strategy exploits both task and data-level parallelism and achieves a 3.5x speedup over single-threaded implementations.

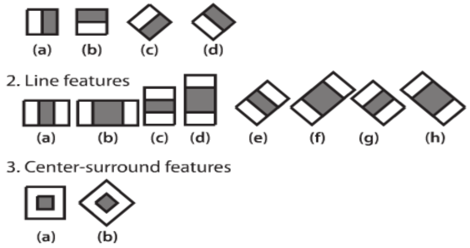


Fig. 3 HAAR Features (Basic & Extended)

The recent work by Jiamin wu *et al.* [13] have proposed new methodology, with the name C^4 for object detection, which has performed very well on face detection too. They have been able to achieve the 20 fps speed and state of the art detection efficiency. They have used and applied the conjecture that contours and signs of comparisons with the neighboring pixels are key information for object detection. The C^4 means, Contour, Cascade Classifier and CENTRIST visual descriptor [14]. Authors have proposed the future scope of work in accelerating the speed of operation of C^4.

They have used new visual descriptor CENTRIST [14] which has been found to be suitable match for contour based object detection. CENTRIST visual descriptor encodes the signs of neighboring comparisons. It has been derived from Census Transform (CT) which was originally designed to establish correspondence between neighboring patches [15]. Please see Fig. 4, it shows the calculation of CT value for the center pixel. The CT image C of an input image I is generated by replacing a pixel with its CT value. The CENTRIST descriptor is a histogram with 256 bins, which is a histogram of these CT values in an entire image or a rectangular region in an image.

$$\begin{array}{c|c|c} 32 & 64 & 96 \\ \hline 32 & \mathbf{64} & 96 \\ \hline 32 & 32 & 96 \end{array} \Rightarrow \begin{array}{ccc} 1 & 1 & 0 \\ 1 & & 0 \\ 1 & 1 & 0 \end{array} \Rightarrow (11010110)_2 \Rightarrow \mathrm{CT} = 214$$

Fig. 4 Census Transform

El-Sayed *et al.* [16] have used mean of medians of CbCr color correction approach to enhance the combined SMQT (Successive Mean Quantization transform) features and SNoW (Sparse Network of Winnow) classifiers (SFSC). The proposed method has been found to be more efficient and accurate as compared to original SFSC method.

Viola Jones *et al.,* [17] did seminal work in face detection, implemented state-of-the-art face detector. This algorithm is still continuing to be leader in modern face detection implementations such as in mobile devices, also inbuilt in OPENCV and MATLAB. Key contributions of their work were: Scale invariance, New image representation called integral image to facilitate faster calculation of Haar features, Uses Machine learning, Adaboost learning algorithm for combining the weak classifier to obtain strong classifiers, Cascade of haar classifiers (see Fig. 3) for faster computation. It was 15 times faster than previous work, can be generalized to detect any type of object. There were some limitations such as the training set of negative examples has to be small to make training feasible, the process of creating the detector cascade is based on trial and error process, training process is lengthy (may take few days, depends upon speed of machine), can handle up to limited rotation angle of faces (± 15 degrees in plane, ±45 degrees out of plane), fails in face occlusions and low brightness of faces. See Fig. 5, how Haar features are able to search for various face features such as eyes, nose, lips etc.

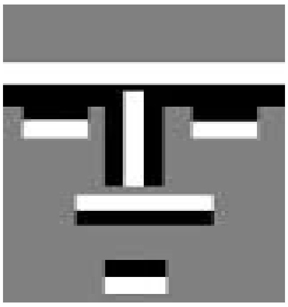


Fig. 4 Applying HAAR Features on Face

Nasrollahi, K. *et al.,* [18] did excellent work on generating good quality frontal face image from low resolution video sequence, used viola and jones [2] face detector, face quality is estimated based upon facial features: sharpness, brightness, resolution, head Pose. They used auto associative memories for the head pose estimation. Further generated high resolution frontal face image using reconstruction and learning based super resolution techniques in cascade.

Bagdanov, A. *et al.* [19] worked on multi-face detection, tracking, facial image quality analysis and face-log generation. They developed multipose face detector, based on Adaboost face detector, used lateral and frontal face detectors. System has been evaluated on 10 hours of realistic surveillance videos, with both quantitative and qualitative analysis. However it reported to have limitations such as the proposed face-logging system is appropriate for situations in which face size is bounded, illumination conditions are consistent with the images used to train the Adaboost detectors in their multipose face detectors.

Chen, Tse-Wei *et al.* [20] combined the spirit of image based face detection and essence of video object segmentation to filter out face candidates. Developed a face scoring technique, using eight scoring functions based on feature extraction technique, used a single layer neural network training system to obtain an optimal linear combination to select high quality faces. The face detector was based on skin color detection and video object segmentation. Scoring functions used eight functions: Skin color coverage, Luminance variation, Circularity measurement, Eye-pixel Histogram, Ratio, Angle, Symmetry and Hair. All eight functions were combined using fuzzy logic to calculate the final score.

Chang-yeon, Jo. *et al.* [21] worked on LBP, face images are divided into M small non-overlapping regions, LBP histogram are extracted from each sub-region, All such histograms are combined together into a single spatially enhanced feature histogram, Extracted feature describes local as well as global shape of the face images, a variant of AdaBoost, Gentle AdaBoost has been used to select the features and train the classifier. Cascade of classifiers is used for enhancing the performance. The developed algorithm has been found to computationally efficient and tested on Mobile platform.

III. Conclusion and Future Scope

There is no doubt that lot of research work has been done in the area of face detection but the goal is still far from achieved: To mimic the human vision of detecting and identifying the human faces. So to meet that goal, still a lot of work has to be done in this area. As per literature survey, following directions for future work in this area are being proposed:

1. The training of Haar features in seminar viola jones' face detector takes a long time, which may be couple of days if used serial processing. There is scope of work to apply the parallel computing to enhance the speed of features training. Till date not much work has addressed the performance
2. Comparisons of various software platforms such as MATLAB, use of GPU in C/C++ environment, use of GPU in MATLAB environment. So there is scope of using optimization work to address the issue of speed of training of features.
3. In the use of volumetric features, there is open research area in: a) Integrating the descriptor with the scanning strategy, b) Setting criteria for selecting the optimal number of frames to encode the descriptor, c) Investigating in using same feature space for face detection & recognition.
4. Use of holistic features for performing various tasks in the process of face extraction from video such as face detection, face quality estimation, face quality enhancement and face recognition instead of using separate feature for each task.
5. Using motion information in creating face-logs from the video.

References

[1] C. Zhang and Z. Zhang, “A survey of recent advances in face detection,” Tech. rep., Microsoft Research, Tech. Rep., 2010.

[2] H.M.M.R. Hiremath, P.S., “Face detection and tracking in video sequence using fuzzy geometric face model and motion estimation,” International Journal of Computer Applications, Vol. 15, No. 58, pp. 12–16, 2012.

[3] H. Seyedarabi, S.M. Bakhshmand, and S. Khanmohammadi, “Multipose head tracking using colour and edge features fuzzy aggregation for driver assistant system,” in Signal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on. IEEE, 2009, pp. 385–390.

[4] B. Jun, I. Choi, and D. Kim, “Local transform features and hybridization for accurate face and human detection,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 35, No. 6, pp. 1423–1436, 2013.

[5] S. Bharadwaj, M. Vatsa, and R. Singh, “Can holistic representations be used for face biometric quality assessment?” in Image Processing (ICIP), 2013 20th IEEE International Conference on. IEEE, 2013, pp. 2792–2796.

[6] Y. Martinez-Diaz, H. Mendez-Vazquez, N. Hern´andez, and E. GarcIa-Reyes, “Improving faces/ non-faces discrimination in video sequences by using a local spatio-temporal representation,” in Biometrics (ICB), 2013 International Conference on. IEEE, 2013, pp. 1–5.

[7] B. Bilgic, B.K. Horn, and I. Masaki, “Efficient integral image computation on the gpu,” in Intelligent Vehicles Symposium (IV), 2010 IEEE. IEEE, 2010, pp. 528–533.

[8] H. Jia, Y. Zhang, W. Wang, and J. Xu, “Accelerating viola-jones facce detection algorithm on gpus,” in High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICESS), 2012 IEEE 14th International Conference on. IEEE, 2012, pp. 396–403.

[9] J. Kong and Y. Deng, “Gpu accelerated face detection,” in Intelligent Control and Information Processing (ICICIP), 2010 International Conference on. IEEE, 2010, pp. 584–588.

[10] E. Li, B. Wang, L. Yang, Y.T. Peng, Y. Du, Y. Zhang, and Y.J. Chiu, “Gpu and cpu cooperative accelaration for face detection on modern processors,” in Multimedia and Expo (ICME), 2012 IEEE International Conference on. IEEE, 2012, pp. 769–775.

[11] D. Oro, C. Fern’ndez, C. Segura, X. Martorell, and J. Hernando, “Accelerating boosting-based face detection on gpus,” in Parallel Processing (ICPP), 2012 41st International Conference on. IEEE, 2012, pp. 309–318.

[12] G. Wei and C. Ming, “The face detection system based on gpu+ cpu desktop cluster,” in Multimedia Technology (ICMT), 2011 International Conference on. IEEE, 2011, pp. 3735–3738.

[13] J. Wu, N. Liu, C. Geyer, and J.M. Rehg, “C¡ sup¿ 4¡/sup¿: A real-time object detection framework,” 2013.

[14] J. Wu and J.M. Rehg, “Centrist: A visual descriptor for scene categorization,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 33, No. 8, pp. 1489–1501, 2011.

[15] R. Zabih and J. Woodfill, “Non-parametric local transforms for computing visual correspondence,” in Computer Vision XECCV’ 94. Springer, 1994, pp. 151–158.

[16] M.A. El-Sayed and N.G. Ahmed, “Enhanced face detection technique based on color correction approach and smqt features,” Journal of Software Engineering and Applications, Vol. 6, pp. 519, 2013.

[17] P. Viola and M.J. Jones, “Robust real-time face detection,” International Journal of Computer Vision, Vol. 57, No. 2, pp. 137–154, 2004.

[18] K. Nasrollahi and T.B. Moeslund, “Extracting a good quality frontal face image from a low-resolution video sequence,” Circuits and Systems for Video Technology, IEEE Transactions on, Vol. 21, No. 10, pp. 1353–1362, 2011.

[19] A. Bagdanov, A. Del Bimbo, F. Dini, G. Lisanti, and I. Masi, “Compact and efficient posterity logging of face imagery for video surveillance,” IEEE Multimedia, Vol. 19, No. 4, pp. 48–59, 2012.

[20] T.W. Chen, S.C. Hsu, and S.Y. Chien, “Automatic feature-based face scoring in surveillance systems,” in Multimedia, 2007. ISM 2007. Ninth IEEE International Symposium on. IEEE, 2007, pp. 139–146.

[21] Chang-yeon, “Face detection using lbp features,” 2008.

Audio Steganography Using LSB Edge Detection Algorithm

Navneet Kaur[1] and Sunny Behal[2]
[1,2]*Department of Computer Sc. & Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozepur, India
E-mail: [1]navneet18kaur@gmail.com, [2]sunnybehal@rediffmail.com

***Abstract*—Digital Steganography is used to protect digital content or data such as text, images, audio and videos that have been tampered maliciously. In this paper we maintain the quality of audio and image and to ensure the ownership, we propose a new LSB (least significant bit) using edge detection in digital steganography. We apply a 2-level steganography on images and audio which make the data which may be text or image in more secure form. By using lsb techniques may effects less the image pixel quality and audio sound quality, in this wecan randomly selected edges and embed the text or image by considering image quality, audio quality and audio imperceptibility.**

Keywords:* *Digital Steganography, Audio Steganography, LSB (Least Significant Bit), Performance Evaluation Metrics, Lsb Algorithm

I. Introduction

Digital Steganography is the technique of securing digitized data by hiding it into another piece of data which may be any text, image, audio, and video. The main task of the field of steganography is the storing, hiding, and embedding of secret data in all types of digital data. [1]The main goal of steganography is to communicate securely in a completely undetectable manner [2] such that no one can suspect that it exist some secret information. Unlike cryptography, which secures data by transforming it into another unreadable format, steganography makes data invisible by hiding (or embedding) them in another piece of data [3] [4].

In this type of steganography we can embed secret messages into digital sound in audio steganography. It is more complex process as compare to embedding messages in other media. This steganography method can embed messages in WAV, AU And even MP3 sound files [5]. The audio steganography consists of Carrier or Audio file, Message and Password. Carrier is also known as a cover-file, which conceals the secret information. In steganography model the secret message that the sender sends wants to remain it secret.[6] Message can be of any type may be text, image, audio or any type of file,.in secret stego key which only the receiver knows the corresponding decoding key will be able to extract the message from a cover-file. The cover-file with the secret information is known as a stego-file [7].

Applications of Audio Steganography

1. Confidential communication and secret data storing.
2. Protection of data alteration
3. Access control system for digital content distribution
4. Media Database systems.
5. To improve the quality.

The rest of the paper is organized as section 2 defines the Related Works. Section **3** describes the Performance Evaluation Metrics. Section 4 describes Proposed Architecture. Section 5 describes Proposed Algorithm. Section 6 defines Experimental Results.

II. Related Work

The techniques involved in audio steganography are:

1) Echo Hiding

Echo hiding used to embeds secret data in a audio file by pass an echo into the discrete signal. This technique has advantages of providing a high data transmission rate and robustness when we make comparison of echo hiding to other methods [8] [9].

2) Phase Coding

Phase coding exploits HAS insensitivity to relative phase of different spectral components. In this method we can replace selected phase components from the original sound signal spectrum with hidden information.due to in audibility of information, phase components medication should be kept small [10].

3) Parity Coding

This technique is one of the robust audio steganographic techniques. Instead of breaking a signal into individual samples, it breaks a signal into separate samples sections and embeds each bit of the secret message information from a parity bit. If the of a selected parity bit region does not match the secret message bit to be encoded, the process inverts the LSB of one of the section in the region. Then the sender has many choices for encoding the secret bit [11].

4) Spread Spectrum

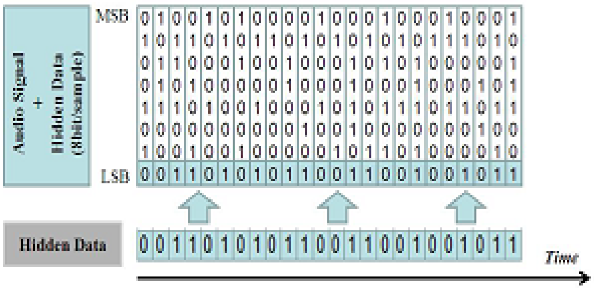
In this technique spread out the encoded information across the available frequencies. This is analogous to a system using an implementation of the LSB coding that randomly spreads the message bits over the entire sound file [12][13].

5) Tone insertion

Tone insertion used on the inaudibility of lower power tones in the presence of significantly higher ones. This method used resist to attacks such as low-pass filtering and bit truncation [14][15].

6) LSB (Least Significant Bit)

LSB in images: It is a simple approach to embedding information in image, in other mainly image manipulation can destroy the hidden information in the image. Due to this by applying LSB to each byte of 24 bit image, 3 bits can be encoded to each pixel or each pixel can be encoded by 3 bytes. Applying LSB technique each byte of 8 bit image only one bit can be encoded into each pixels as each pixel is represented by one byte.



In Audio LSB coding, two least significant bits of a data is replaced with two message bits. If we increase the amount of information encoded will also increase the noise in the sound file. Like, a sound file that was recorded in a bustling subway station would mask low-bit encoding noise[16]In secret message extraction from an LSB encoded audio file, the recipient needs access to the sequence of sample indices used in the embedding process. The length of the secret message to be encoded is smaller than the total number of section in audio file. We also know about how to choose the subset of samples which contain the secret message or information and communicate that decision to the recipient[17]. One trivial it is to start at the beginning of the audio file and perform LSB coding unto message completely embedded, leaving the remaining sections unchanged. But it creates a problem like in the first part of the audio file will have different statistical properties than the second part of the audio file which was not modified. Solution to this problem is to pad the secret message with random bits so that the length of the message is equal to the total number of samples. LSB (Least Significant Bit), this method is one of the important and easiest methods used for data hiding [18]. Traditionally, it is based on embedding each bit from the message in the least significant bit of the cover audio in a deterministic way.

Advantages: more embedding capacity for information and easy to implement or to combine with other hiding methods.

Disadvantage: less robustness to noise addition which reduces its security performance since it becomes vulnerable even to simple attacks [19].

III. PERFORMANCE EVALUATION METRICS

The performance of the watermarked images must be evaluated by using some quality measures such as MSE, SNR, PSNR and BER.

1) The MSE (Mean Square Error): [20] Defined it as Average Squared Difference Between a Reference Image and a Distorted Image. It is Calculated as:

$$MSE = \frac{1}{XY}\sum_{i=1}^{X}\sum_{j=1}^{Y}(c(i,j) - e(i,j))^2$$

X represents the height and Y represents the width of the image c (i,j) and e (i,j) are the respective pixel value of the original image and embedded image. [20]

2) The PSNR (Peak Signal to Noise Ratio): It is a Quality Metric Used to Determine the Degradation in the Embedded Image with Respect to the Host Image or also Defined as Ratio between Maximum Power of a Signal and Power of Distorted Signal [20]. It is Most Easily Defined via the Mean Squared Error (MSE) as:

PSNR = $10log_{10}\frac{L*L}{MSE}$

L denotes the peak signal value of the cover image which is equal to 255 for 8 bit images.

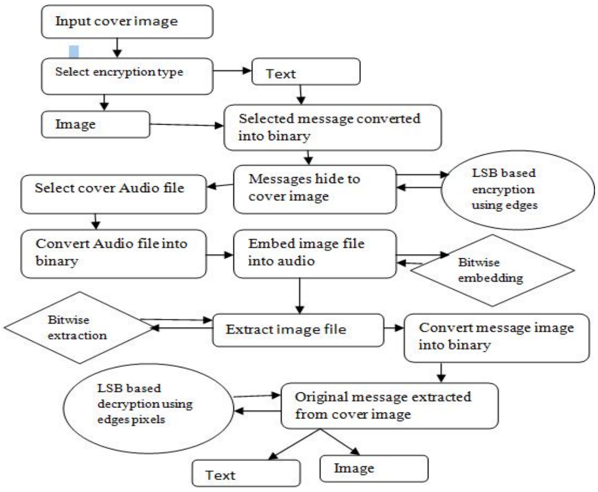

Input cover image
Select encryption type
Text
Image
Selected message converted into binary
Select cover Audio file
Messages hide to cover image
LSB based encryption using edges
Convert Audio file into binary
Embed image file into audio
Bitwise embedding
Bitwise extraction
Extract image file
Convert message image into binary
LSB based decryption using edges pixels
Original message extracted from cover image
Text
Image

Fig. 1 Proposed Architecture

IV. PROPOSED ARCHITECTURE

We used this method to make the security level of steganography more secure against attacks. Purposed method is consisting of 2-Level Security Process. In this first we select any input cover image then select encryption type which may be text or image and then message converted into binary. After conversion message hide to cover image by LSB based encryption using edges. Then select a cover audio file, then convert audio file into binary, then embed image file into audio file by bitwise embedding.

Extraction Process: After embedding we can extract image file from audio file by Bitwise Extraction and then convert into binary. After converting original message extracted from cover image which may be text/image by LSB based decryption using edge pixels.

V. Proposed Algorithm

In Audio Steganography, we use a least significant bit using edge detection.

A. Embedding Algorithm

1. Select an Input cover image.
2. Select Encryption type which may be an image or text.
3. Selected message(text/image) converted into binary.
4. Using least significant bit (LSB) based Encryption using Edges to message hide in cover image.
5. Select an Audio cover file.
6. Convert selected Audio cover file into Binary.
7. Using bitwise embedding to embed Image file into Audio file.

B. Extracting Algorithm

1. Extract image file using bitwise extraction.
2. Convert message image into binary.
3. Original message which may be text or image is extracted from cover image by least significant bit (LSB) based decryption using Edges.

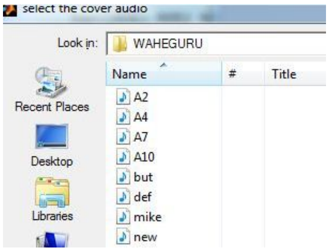
VI. Experimental Results

In Audio steganography, we have made a 2-level steganography. In steganography mainly we can embedding a text in an image or embedding a text in an audio, But in this paper we can modified the methods by combination of two methods to make it 2-level,Firstly we can embed the text message in an image and then embed the encrypted text message in an audio wav file. In embedding first select a cover image of size

Select a cover image and enter the text message which we want to encrypt

Enter a message which we want to encrypt
1111111111111
Encrypted image with hidden message

Select an Audio file and using bitwise embedding to embed encrypted image in audio



Extracting process:
Extract image from audio

Extract Original message from image:
1111111111111

Table I (Experimental Results)

Image	LSB3	Jae Gilyu	First Component Alteration Technique	Improved LSB	LSB Using Edge Detection
PSNR	37.92	38.98	46.11	46.65	68.60

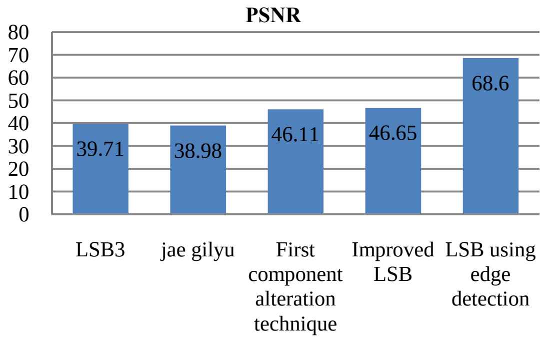


Fig. 1 Comparative Study of Various Methods with Proposed Techniques

VII. CONCLUSION

This paper provides that proposed algorithm is more secure due to 2-level steganography which gives the robustness and good quality of images or audio. In future work focus on size image with respect to time, because as we increase the size of message, increases the size of cover image as well as size of audio wav file increase which consumes more time.

ACKNOWLEDGMENT

This is to express my sincere gratitude to Mr. Sunny Behal, Assistant Professor, Department of Computer Science & Engineering, SBS State Technical Campus, Ferozepur (Punjab), India, for sparking in me the enthusiasm and initiative to discover and learn. I am truly thankful to him for guiding me through the entire paper and being my mentor and guide in this learning curve.

REFERENCES

[1] Artz, Donovan. Digital steganography: hiding data within data internet computing, IEEE 5.3 (2001): 75–80.

[2] Amin, Muhalim Mohamed, *et al.,*. Information hiding using steaganography. Telecommunication Technology 2003. NCTT Proceedings, 4th National conference on IEEE, 2003.

[3] Amin, Shashikala Channalli and Ajay Jadhav, “ Steganography An Art of Hiding Data”, International Journal on Computer Science and Engineering, IJCSE Vol. 1, No. 3, 2009.

[4] Shashkila channalli, Ajay jadhav, “Steganography an art of hiding data” International journal on Computer science and engineering Vol. 1(3), 2009, 137–141.

[5] Prof. Samir Kumar, Bandyopadhyay Barnali, Gupta Banik,” LSB Modification and Phase encoding Technique of Audio Steganography Revisited”. Vol. 1. (4) IJARCCE 2012.

[6] Dalal N. Hmood, Khamael A. Khudhiar and Mohammad S. Altaei (2012). A New Steganographic Method for Embedded Image In Audio File. International Journal of Computer Science and Security (IJCSS) 6(2): pp. 135–141.

[7] Chandrakar, Pooja, Minu Choudhary, and Chandrakant Badgaiyan. "Enhancement in Security of LSB based Audio Steganography using Multiple Files." International Journal of Computer Applications 73 (2013).

[8] HS, Anupama. "Information Hiding Using Audio Steganography A Survey." International Journal of Multimedia & Its Applications 3.3 (2011).

[9] Mat Kiah, M.L., *et al.,*. "A review of audio based steganography and digital watermarking." International Journal of Physical Sciences 6.16 (2011): 3837–3850.

[10] Samir K Bandyopadhyay, Debnath Bhattacharyya, Debashis Ganguly, Swarnendu Mukherjee and Poulami Das, “A Tutorial Review on Steganography”.

[11] Jenkins, Neil, and Jean Everson Martina "Steganography in audio." University of Cambridge CST Part II Dissertation (2009).

[12] Malviya, Swati, Manish Saxena, and Dr Anubhuti Khare. "Audio Steganography by Different Methods". International Journal of Emerging Technology and Advanced Engineering.

[13] L.M. Marvel, C.G. Boncelet Jr., C.T. Retter, Spread spectrum image steganography, IEEE Trans. Image Process. 8 (8) (1999) 1075–1083

[14] Pramatha Nath Basu, Tanmay Bhowmik,’On Embedding of Text in Audio–A case of Steganography’ International Conference on Recent Trends in Information, Telecommunication and Computing.

[15] Nosrati, Masoud, Ronak Karimi, and Mehdi Hariri. "Audio Steganography: A Survey on Recent Approaches." World Applied Programming 2.3 (2012): 202–205.

[16] Kumar, H.; Anuradha "Enhanced LSB technique for audio steganography". Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference o*n,* On page(s): 1–4.

[17] xiaolong Li, Bin yang, Daofong Cheng and Tieyong Zeng “A generalization of LSB matching”, IEEE signal processing Letters, Vol. 16, No. 2, Feb-2009.

[18] Singh, Pradeep Kumar, Hitesh Singh, and Kriti Saroha. “A survey on Steganography in Audio.” National Conference on Computing for Nation Development, India com. 2009.

[19] Nitin jain,Sachin mesh ram and Shikhar dubey, “Image steganography using LSB and EDGE detection techniques”, International journal of soft computing and engineering, ISSN: 2231–2307, Vol. 2, Issue: 3.

High-Speed, Long-Reach Bidirectional OFDM-DWDM-EPON Access System

Shivani Sharma[1] and Vishal Sharma[2]
[1,2]Department of Electronics & Comm. Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]shivanishrm91@gmail.com, [2]er_vishusharma@yahoo.com

***Abstract*—For the increasing demand for capacity of optical transmission systems, Ethernet-PON comes out to be a most prominent and assuring solution in todays world. The deployment of Ethernet-PON provides the evolution path to ever higher bandwidths. With OFDM, system achieves the high bit rate and higher bandwidth over other modulation formats. In this paper, we demonstrate a Ethernet PON based on orthogonal frequency division multiplexing employing OSSB modulation scheme. The simulated demonstration directs the transmission of single channel of 10 Gbps data rate to realize the proposed system and achieves the fiber link of 30km in downstream direction (OLT to ONU). The system performance can be witnessed by measuring the SNR and constellation diagram.**

Keywords: OFDM (Orthogonal Frequency Division Multiplexing), OSSB(Orthogonal Single Side Band), EPON(Ethernet Passive Optical Network)

I. INTRODUCTION

Now a days**,** multimedia applications such as internet protocol television (IPTV) and high definition (HD) video continue to conflagrate the development in bandwidth demand in multi-user access networks. An increase in the number of users are putting pressure on the comm-unication system vendors to offer higher data rates. The passive optical networks are the most significant class of fiber access systems to address the upcoming issues related bandwidth and data rates [1]. The PON technology proves to be very efficient solution to cope with the problems of broadband access networks. Also, PON technology has been confirmed the transmission of triple play servies such as voice, vedio and data in downstream link [2]. EPON systems provides better performance on different wavelengths in both upstream and downstream direction. Recently, an asymmetric GPON originates with the capability of providing data rate of 2.5Gbps in downstream and 1.25 Gbps in upstream channels, simultaneously [3]. In EPON systems, data packets are broadcasted to multiple ONUs in downstream direction and the intended ONU extracts the data packets that are ment for them. Here, EPON simply acts as a point-to-multipoint network and in upstream link, EPON acts as multipoint-to point network. Furthermore, only EPON nodes with privileged traffic can be WDM-upgraded by using either fixed or tunable transceivers [4]. However, there was a problem of chromatic dispersion which significantly limits the transmission distance [5]. As the data transmission speed of communication system goes to increase, a corresponding decay in time for each transmission occurs. Thus, the Intersymbol Interference (ISI) becomes a severe limitation on the high data rate communication[6]. To cope with this problem, OFDM technology along with PON comes into existence. OFDM yields high transmission rate and preferred spectrum utilization by making use of M-ary modulaton techniques such as QAM, PSK [7]. An OFDM-EPON allows flexible assignment of 4-QAM at 250 Mbps and 16-QAM LAN traffic at 500 Mbps bandwidth by allocating different number of subcarriers [8]. It is highly flexible in terms of supporting multiple granularities of bandwidth through high efficient digital modulation and dynamic resource management. The sharing of that bandwidth to all users is possible which is allocated to the sub-carriers through TDM mode [9]. The higher level of quadrature amplitude modulation (QAM) helps to increase the aggregated data rate while preserving the data bandwidth as it was i.e by making use of same number of OFDM subcarriers. The receiver sensitivities tends to increase on account of falling decision margins in the EVM calculations[13]. The participation of orthogonal frequency division multiple access (OFDMA) technology for the LAN traffic transmission does not requires any change in existing EPON architecture. A single receiver at optical network unit (ONU) can detect both LAN and EPON downstream traffic, that makes the system economically good. Also, flexible assignment of LAN traffic bandwidth is analyzed by making use of different modulation formats as well as by assigning different number of subcarriers [14].Although optical fiber bandwidth is larger but the optical devices exhibits the limited bandwidth and the devices providing larger bandwidth gives rise to increase in cost. Hence, the bandwidth efficiency is the major issue. The OSSB transmission is the feature of modulation system that prove to be brilliant solution to increase the bandwidth efficiency by factor of two [10]. OFDM PON not only solved the problem of optical access network speed by improving the transmission speed, even it proved very convenient, low cost and flexible upgrade in the technology. This OFDM bsased optical access technology now become the tumid research hotspot [12]. In this work, we have demonstrated an OFDM based EPON system at 1550nm to measure the system performance at high data rate which is not elaborated earlier. The proposed system is investigated for successful transmission of downstream channel over the SSMF at high data rate and fiber link between ONU and

OLT. The system is optimized at acceptable SNR of 15 dBm. This paper is organized as follows: Section I briefly describes need and year to year development in EPON system. Section II deals with the portrayal of simulated OFDM based EPON model and to measure the results. The section III puts light on the measured results in section II.Simulated OFDM based DWDM-OSSB-EPON model and discussion on simulated findings. The section III puts light on the measured results in section II.

II. Model Description & Result Discussion

The proposed customer access EPON system using OFDM is shown in Fig. 1. In our proposed OFDM-PON system, Quadrature Amplitude Modulated(QAM) data signals are generated using 4QAM sequence generator consisting of 2 bit per symbol. This QAM data signals are then modulated by OFDM modulator which uses 512 subcarriers and FFT size of 1024 to generate OFDM analog data signals. These signals are then QAM modulated at 7.5 GHz frequency. This QAM-OFDM treated analog data signals are then modulated by means of LiNbo$_3$-dual electrode MZM modulator, phase of that signals gets shifted by phase shifter and an optical source at 1550 nm to generate OSSB signals. These signals are then transmitted over SSMF fiber without using any active device in between OLT and ONU. All design parameters are taken into consideration as IEEE 802.3 ah standard [11]. The SMF fiber parameters are chosen as nonlinear reflective index coefficient = 2.6 × 10^{-20} m^2/W; effective area, A$_{eff}$ = 80 m^2; attenuation = 0.2 dB/km; dispersion = 17 ps/nm/km; and dispersion slope = 0.075 ps/nm^2/km. The system is simulated using Optisys^{TM} software.

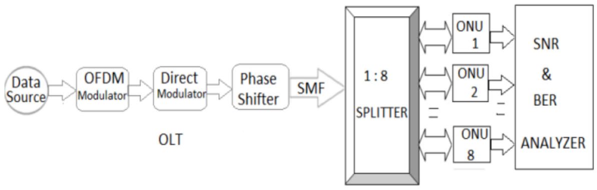
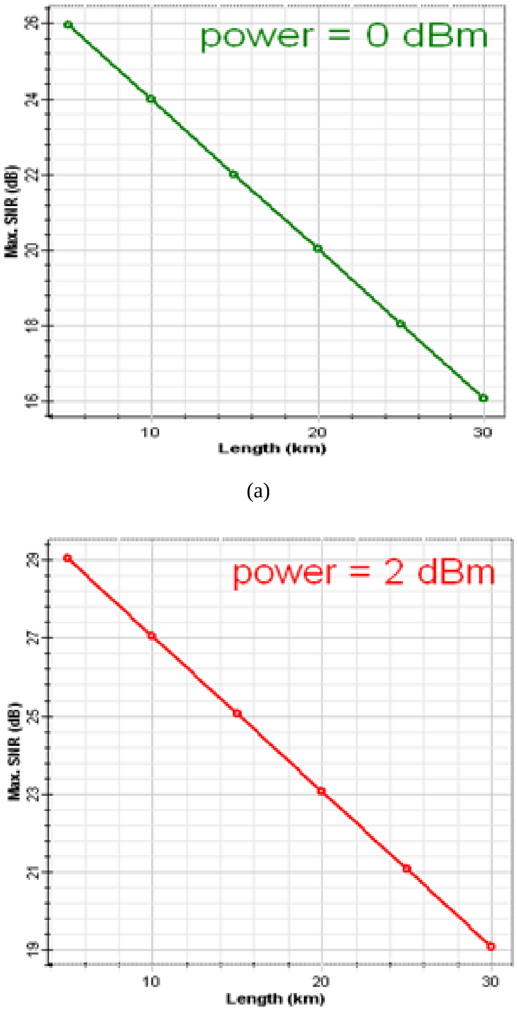


Fig. 1 Block Diagram of Downstream OFDM-OSSB-EPON System

The PIN photo-detector for converting the optical signal into electrical signal with sample rate of 40 GHz; responsivity of 1 A/W; dark current of 10nA; OFDM demodulator with 512 sub-carriers, position array 256 and number of FFT points 1024 are used at the receiver end for electrical transmission. Different spectrum analyzers such as SNR analyzer to measure the signal to noise ratio and BER analyzers for bit error rate measurement are connected after 3R regenerator at the receiver side to investigate the observations.An optical power meter is also connected after the passive splitter to check signal strength at each node.

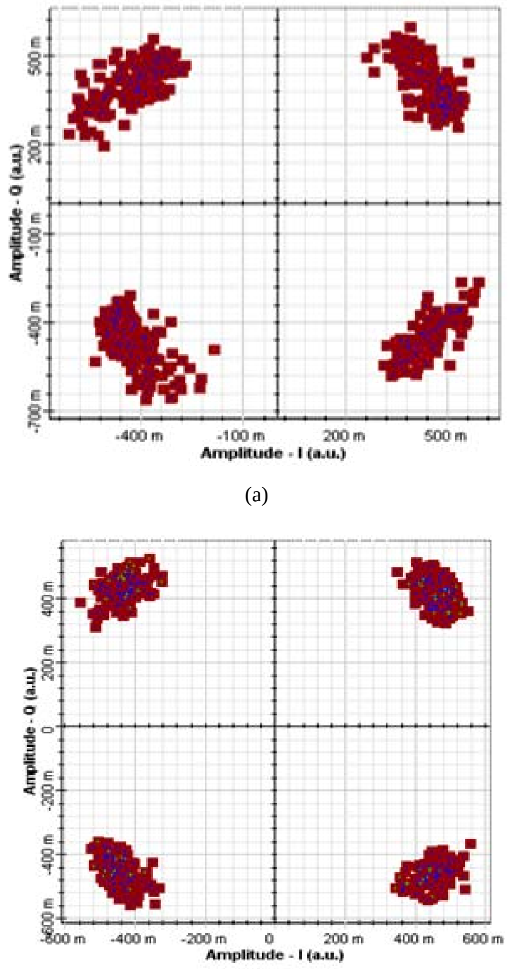
For the generation of OFDM treated single side band signals, the input channel is applied to both the electrode of MZM modulator suchlike that the input signal is given directly at one electrode and at another electrode with 90^0 phase shift. This OFDM treated OSSB modulated channel at 1550 nm of 10 Gbps data rate is transmitted over SSMF simultaneously at distance of 30km to realize OFDM- based OSSB-PON access system.The OSSB modulation technique is the most efficient technique for transmitting baseband digital data with minimum fiber dispersion over long distances. There is no any active component such as amplifier is used between optical line terminal (OLT) and optical distribution network (ODN).Then at the optical distribution end (ODN) a Passive splitter which acts as a Hub will split the single channel data into eight different ONUs. The data packets will get received by only intended ONU. Each onu can recognize its own data packets depending upon the MAC address. Then at the receiver end, the photo detector will convert the optical transmitted signal into the electrical signal for further use. The analysis of the demonstrated system are teken by using different analyzers.



(a)

(b)

Fig. 2 Measured SNR vs. Channel Length at Laser Power (a) 0dBm (b) 2dBm, at 10 Gbps



(a)

(b)

Fig.3 Measured Constellation Diagram at Data Rate 10 Gbps at Power (a) 0 dBm (b) 2 dBm

As demonstrated, the measured SNR realization for downstream EPON system is shown in figure 2 at different power. The system is capable of achieving the fiber length of about 30 km with split ratio of 8 at 1550 nm wavelength. On comparing the results obtained in Fig (a) & (b), an improvement of about 3 dBm is acheived in SNR at 2 dBm power as shown in Fig (b) than SNR at 0 dBm as shown in fig(a) over an optical span of 30 km.it is noticed that the decay in SNR with the incrase in channel length is more at 0 dBm. The measurements of constellation diagram are shown in fig.3. The signal shows the great improvement in terms of quality at 2dBm than at 0 dBm. Here, the blue points shows the noise that comes from the laser diode and the red points indicates the signal. So it is indicated that the OFDM based EPON system provides better system performance with high split ratio with low BER.

III. Conclusion

We have proposed and demonstrated a multi-carrier multiplexing i.e OFDM based OSSB-EPON system. The system performs much better when laser power is increased. High split ratio over a long optical span with low BER is achieved by single channel transmission at 10Gbps. Hence, OSSB scheme is very effective in providing bandwidth efficiency and low chromatic dispersion over long haul communication. From this results, we have concluded that system shows the SNR within the acceptable limits even at low power. Accordingly, OFDM based system is recommended to accomplish better optical span with high split ratio at SNR within the acceptable limits.

References

[1] Frank Effenberger, David Cleary, " An introduction to PON technologies", "IEEE communication", pp. 0163-6804, 2007.

[2] J.O. Farmer, "Delivering video, voice and data to consumers via an all fiber network, " consumer electronics, in: Digest of Technical Papers, ICCE International Conference, 2002, pp. 158–159, http://dx.doi.org/10.1109/ICCE. 2002.1013971

[3]] Y.M. Lin and P.L. Tien, "Next-generation OFDMA-based passive optical net-work architecture supporting radio-over-fiber, " IEEE journal of Select Areas Communication, Volume28, issue 6, August 2010, pp.791-799.

[4]] M.P. McGarry, M. Maier and M. Reisslein, "WDM Ethernet passive optical networks (EPONs), " IEEE Communication Magazine, Volume 44, no. 2, pp. 15-22, Feb. 2006

[5] H.Schmuck, "Comparison of optical millimeter-wave system concepts with regard to chromatic dispersion", "Elsevier journal", pp.1848-1849, Oct 1995

[6] Neha pathal Mtech scholar, Shriam Institute of Technology, "OFDM (orthogonal frequency Division Multiplexing)simulation using Matlab", International journal of Engineering & Technology(IJERT), Vol.1, issue 6, August 2012, ISSN:2278-0181

[7] Fahad.Almasoudi, Khaled alatawi, "Study of OFDM Technique on ROF passive optical network", "Optics and photonics journal", vol.3, pp.217-224, 2014

[8] Lei Deng, Ying Zhao, Xiambin Yu, Valeria Arlunno, Robert Borkowski, Deming Liu, Idelfonso Tafur Monroy, "Experimental Demonstration of an improved EPON architecture using OFDMA for bandwidth scalable LAN emulation" Optik, Volume 17, Issue6, December2011, pp.554-557; www.elsevier.de/ijleo

[9] Qi Shao, Chaoqin Gan, Ruixue Wang, Qiongling Shi, "Dynamic bandwidth allocation of both upstream and downstream in Orthogonal Frequency Division Multiplexing Passive Optical Network, " Optik- International Journal for Light and Electron Optics, Volume 124, Issue 18, September 2013, pp. 3476-3479; www.elsevier.de/ijleo

[10] Morant Maria, Llorente Roberto, Hauden Jerome, Quinlan Terence, "Dual-drive LiNbo3 inferometric Mach-Zehnder architecture with exyended linear regime for high peak-to-average OFDM-based communication systems", Opt Express; vol 19, 2011

[11] IEEE P802.3ah task force www.elsevier.de/ijleo

[12] BIN Wu, Zhenrong Zhang, " Applications of Orthogonal frequency division multiplexing in optical access network", "International conference on Enviornmental science and information application technology", vol 10, pp. 1199-1204, 2011

[13] Chow Chin-Wai, Chen-Hung, " WDM extended reach passive optical networks using OFDM-QAM", "OSA journal, in OPTICAL express", vol.16, 2008.

[14] Qi shao, Caoqin Gan, "Dynamic bandwidth allocation of both upstream and downstream in orthogonal division multiplexing", in elsevier journal Optik, vol.124, pp.3476-3479, 2013.

Hybrid Improved Max Min Ant Algorithm for Load Balancing in Cloud

Rajwinder Kaur[1] and Navtej Ghumman[2]
[1,2]Department of Computer Sc. & Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, India
E-mail: [1]rajwindersandhu40@gmail.com, [2]navtejghumman@yahoo.com

***Abstract*—Cloud computing provide unlimited resources to the customers. Cloud also provides many services to the end users. Users can access these resources directly through internet. Users pay only for those resources which they use. In cloud computing environment load balancing is very important concept. Modified max –min used the concept of original max -min. Modified Max-min is based on the execution time not on complete time as a selection basis. This paper proposes hybrid Improved max min Ant Optimization algorithm. The main motive of our work is to balance the total system load .We try to minimizing the total makespan. We simulated results using the Cloud Sim toolkit. Results show the comparison between improved max min and new hybrid improved Max-min ant approach. It mainly focuses on total processing time and processing cost.**

Keywords: Cloud Computing, Load Balancing Algorithms, Improved Max-Min Algorithm, Hybrid Improved Max Min ant Algorithm

I. INTRODUCTION

Cloud Computing is a new technology. Cloud service providers provide many services to end user. All the services are providing on internet with lower cost. Also, there are some applications are provide which users can used through internet. Applications are provide on lower cost in the cloud environment. Users pay only for those resources which they used.[1] They need to pay as much they used. In cloud environment resources are increase quickly. So load balancing is a main problem in cloud computing. In cloud many tasks are executed on available resources at same time. Load balancing is required for proper utilization of all the resources and for better response. Many algorithms are implemented for load balancing in cloud computing. All the algorithms work in different ways. In Cloud computing datacenters are used to collect all the resources; that all resources are shared by many users through the internet.

Cloud computing is a internet based computing service that is provided by infrastructure providers, they provide the services on the users demand. In cloud computing Quality of Service and Load Balance (LB) is very important terms. To improve the performance of resources and for better results many effective task scheduling algorithm are used. Load balancing algorithm is responsible for managing all the jobs that are submitted to cloud environment. Jobs are assigned onto available all resources in such a way that the total response time should minimized. [2][3]. These algorithms also help in manage the makespan. Many of these algorithms try to minimize the total tasks completion time. Max-min, Min-min and RASA are the three well algorithms which used for load controlling. In all these algorithms first calculate the completion and execution time of each task on each available resource in cloud computing. RASA is a combination of other two algorithms. In the RASA algorithm, first completion time of each task is calculated on all available resources then Max min and Min-min algorithms are applied on those tasks. RASA used these algorithms in such a way to take advantage of algorithms and avoids the drawbacks of both algorithms. In Max-min algorithm large tasks to be executed firstly, which means small task delays for long time because they executed after the completion of long tasks. On the other hand, Min-min is executing smaller tasks firstly then large ones that mean long time tasks face delays. [4] [5]

But we used improved max min algorithm. The main idea of an improved Max-min algorithm is that assign task with maximum execution time to resource with minimum complete time at place of original Max-min assign task with maximum completion time to resource with minimum execution time. First we calculate improved max min algorithm result and then we apply improved max min and Ant colony approach together for better results.

A. Metrices for Load Balancing Algorithms

There are many matrices are used in load balancing .These all matrices are helps in measure the performance of algorithms:

1. *Scalability:* It means algorithm is able to perform when numbers of tasked are increased quickly in cloud environment. Any algorithm is good when this metric should be improved as compared to other algorithms.
2. *Resource Utilizations:* It means all the resources are used properly. Better resource utilization It should be better for an efficient load balancing algorithm.
3. *Fault Tolerance:* It means recovery from all types of failure. The load balancing should be a good fault tolerant technique. The main faults which occur are like node failure.
4. *Response Time:* It is the time that is taken by a particular algorithm to respond a task. A good algorithm takes minimum time to respond a task.

5. *Overhead Associated:* It define total amount of overhead involved when implementing a load balancing algorithm in cloud computing environment. Overhead occurs due to the inter process communication between the tasks. If this minimized that load balancing algorithm work properly.

B. *Dynamic Load Balancing*

Types of Load Balancing Algorithms:

a. *Distributed algorithms:* In this all nodes of algorithm execute together, they divide load balancing work between them properly. All the nodes are interacting with each other in two ways: one is cooperative and other is non-cooperative. In the cooperative all the nodes work together but side-by-side to achieve a common goal. .All nodes work together to improve the response time of the system. In Non-Cooperative, all nodes works independently to achieve a common goal for better response time for all tasks.

b. *Non-distributed algorithms:* In this load balancing is done by one or many resources. Non-distributed dynamic load balancing algorithms have two types: one is centralized approach and other is semi-distributed approach. In centralized approach, the load balancing algorithm is executed only by a single node. In this only central node is responsible for all operation and controlling. In semi distributed approach, all nodes of the system are divided into clusters. Then load balancing is done in centralized form in the clusters. A central node controls all the operations of load balancing. That central node is select by choice in each cluster during load balancing operations

II. Dynamic Combination of Improved Max Min and Ant Algorithm

A. *Improved Max Min Algorithm*

In Max-min algorithm large tasks have highest priority and smaller tasks have lower priority. It means, when we have one long task, then Max-min algorithm could execute many short tasks concurrently while executing large task. The make span is calculated in this by the execution of long task .It would be similar to the Min-min make span. [6] [7]

We try to minimize waiting time of short jobs through assigning large tasks to be executed by slower resources. On the other hand execute small tasks concurrently on fastest resource to finish large number of tasks during finalizing at least one large task on slower resource.

Where meta-tasks contain tasks with different completion and execution time, we proposed a new Max-min algorithm that helps in increasing the efficiency of max min algorithm. That is known as improved max min load balancing algorithm.

Algorithm 1: Improved max min load balancing algorith

```
1. For all submitted tasks in meta-task; $T_i$
2. For all resources; $R_j$
3. Cij = Eij + rj
4. While meta-task is not empty
5. Find task $T_k$ costs *maximum execution time.*
6. Assign $T_k$ to the resource $R_j$ which
   gives *minimum completion time.*
7. Remove $T_k$ from meta-tasks set
8. Update $r_j$ for selected $R_j$
9. Update Cij for all j
```

Improved max min increases the chances of execution of tasks on resources.

Algorithm 1 represents improved max min approach. It defined how it works. Max-min algorithm is followed to implement improved Max-min. Load balancing algorithms enhances performance in distributed systems. Sometimes these algorithms not help in better makespan.

There are many existing load balancing algorithms in cloud computing which used for load balancing. Some new algorithms are also implemented from existing algorithms, this will helps to researchers to carry out further work in this area. We combine improved max min and ant algorithm as hybrid approach. Improved max min work in different way from original max min algorithm. [8] [9] [10]

In the original max min "Select task with max execution time then assign that task to that resource which take min completion time” but in improved max min we "Select task with max completion time and then assign resource which take min execution time”

B. *Ant Colony Algorithm*

Ant Colony Optimization (ACO) algorithm is inspired from real ant colonies and it work based upon their actual behavior. Ants are live in colonies. They are work for the survival of colony. Ants always travel from their nest and food sources when they searching for food. In the initial stage ants explore the area surrounding their nest in a random way. While moving from one place to another ants deposit special substances called pheromones. Ants can smell pheromones. When choosing their way for food. During the return trip they again follow the pheromones. Value of pheromones depends on the quantity and quality of the food which ant search. The pheromone trails will guide other ants to the food source.

Algorithm 2 represents the ant colony algorithm approach. Basic method of ACO work is as follow: Ant Colony Optimization (ACO) algorithm is inspired from real ant colonies and it work based upon their actual behaviour.

Algorithm 2: Improved max min load balancing algorith

```
begin
Initialize the pheromone
While stopping criterion not satisfied do Position each ant in a starting VM
while (stopping when every ant has build a solution) do
for each ant do
Chose VM for next task by pheromone trail intensity
end for
end while
Update the pheromone
end while
end
```

Ants are live in colonies. They are work for the survival of colony. Ants always travel from their nest and food sources when they searching for food. In the initial stage ants explore the area surrounding their nest in a random way. In the way they drop pheromones, so other ant follow the same shortest root.

III. Basic Principle of Combination of Improved Max Min and Ant Algorithm

Improved max min provides optimal solution during the preliminary stage, but it will significantly reduce after some time. [11] [12] [13] [14]

However, during the starting stage of ant algorithm, the searching speed is very slow for lacking of pheromones, and then after pheromones reach a certain degree, the speed of optimal solution improves quickly.

The basic principle of dynamic combination between max min algorithm and ant colony algorithm is that we can utilize max min algorithm to get advantages in initial stage and then obtain the optimal solution by ant algorithm in last stages. .

IV. Stimulate Result

Cloud computing simulation platform Cloud sim is used for implementation. It was designed by Australia Melbourne University. In Cloud Sim we used Java as discrete event simulation engine.

Figure 1 represents that how the tasks are allocated. It known as task allocation policies.

Cloud Sim helps in simulation of IaaS (Infrastructure as a Service), PaaS (Platform as a Service) and SaaS (Software as a Service). It provides all the basic components such as Hosts, Virtual Machines, and applications that help in to model the three types of services.

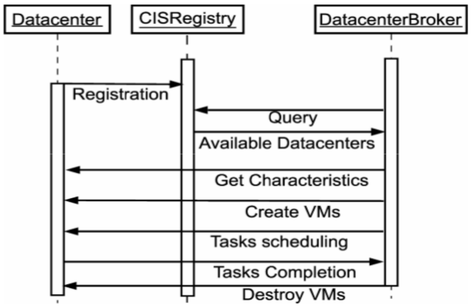


Fig. 1 Model of Tasks Allocation Policies

Figure 2 shows cloudsim workflow model. It defined how cloudsim work in actual. It defined about about different layers.

CIS: It is known as Cloud Information Service. It provides database level match-making services. CIS maps user requests to best suitable cloud providers.

Data centre Broker: it help for mediating between users and service providers. It depending on QoS requirements and the broker deploys service tasks.

Vm Scheduler: Vm scheduler represents the policies. There are two types of policies. One is known as space-shared and second is time-shared. These are required for allocating processing power to all VMs.

Vm Allocation Policy: Vm Allocation Policy is used to select available host in a datacenter. This is mainly focus on the memory, storage, and availability requirement for a VM deployment.

We calculate the result with improved max min that we define as old algorithm and with our hybrid algorithm that is combination of improved max min and ant algorithm. We calculate all the result in cloudsim. We attach the load of planet lab. We calculate result with different number of cloudlets. [15] [16] [17]

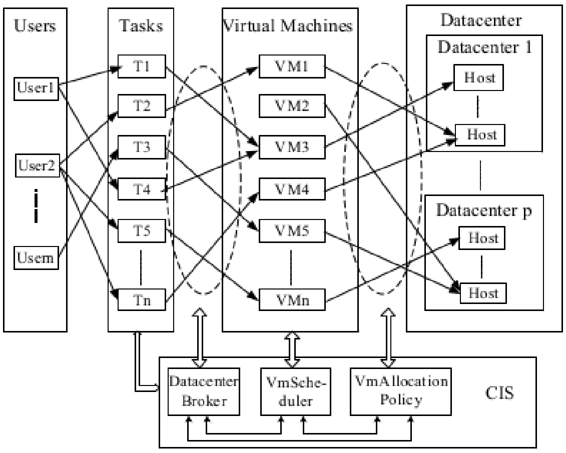


Fig. 2 Cloud Sim Workflow Model

We fix the virtual machine parameter values at the starting.VM parameters are MIPS (millions of instructions per second) and BW (bandwidth). We assign the ID to every VM.

In max min if can't execute tasks concurrently, makes pan become large. To overcome such limitations of improved Max-Min algorithm, a new modification is applied for Max-min scheduling algorithm. To improve the total processing time and cost we applied ant approach on this with improved max min algorithm. Which known as hybrid algorithm. Our approach improve total cost and time factor. This study is only concerned with the number of the resources and the tasks.[18]

Figure 3 shows total processing cost for 1000 tasks (cloudlets).

Figure 4 represents total processing cost for 1000 tasks (cloudlets).We compare the result of old (improved max min algorithm) and hybrid algorithm (combination of improved max min algorithm and ant algorithm)

Figure 5 shows the comparison for different number of tasks. We compare total processing cost

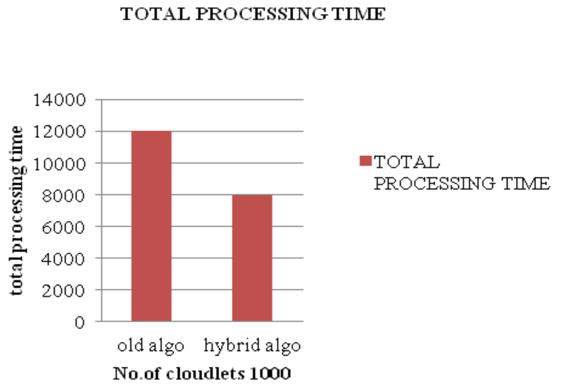


Fig. 3 Total Processing Time for 1000 Cloudlets

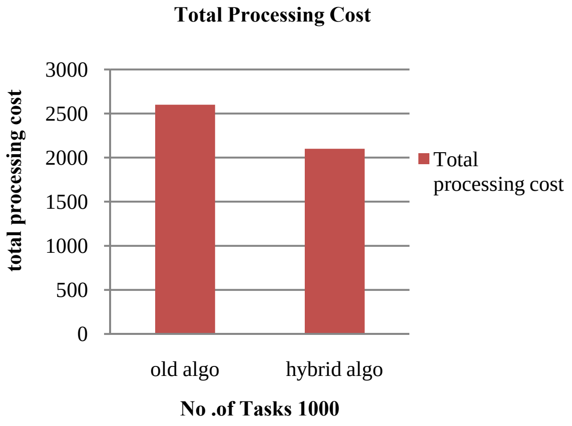


Fig. 4 Total Processing Cost for 1000 Tasks

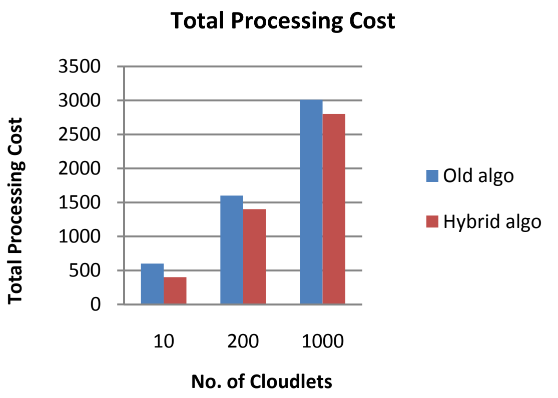


Fig. 5 Comparison for Total Processing Cost with Different No. of Tasks

V. CONCLUSION

We have presented a hybrid improved max min ant approach which is performed by cloud Sim. This helps in better load balancing. It provide better processing time and total processing cost as compared existing algorithms .Load balancing is used to obtain better resource utilization and performance .This study is concentrated only on tasks and resources.

ACKNOWLEDGMENT

The Success of this research work would have been uncertain without the help and guidance of a dedicated group of people in our institute SBS State Technical Campus, Ferozepur India. I would like to express our true and sincere acknowledgements as the appreciation for their contributions, encouragement and support.

REFERENCES

[1] Sahu, Yatendra and Pateriya, RK, Cloud Computing Overview with Load Balancing Techniques., International Journal of Computer Applications, Year 2013, Volume 65

[2] Mell, Peter and Grance, Tim, “The NIST definition of cloud computing”, National Institute of Standards and Technology, 2009, vol53, pages50, Mell2009

[3] Chaudhari, Anand and Kapadia, Anushka, “ Load Balancing Algorithm for Azure Virtualization with Specialized VM”, algorithms, vol 1, pages 2, Chaudhari

[4] Nayandeep Sran, Navdeep Kaur, “Comparative Analysis of Existing Load Balancing Techniques in Cloud Computing ”, vol 2, jan 2013

[5] Chaczko, Zenon and Mahadevan, Venkatesh and Aslanzadeh, Shahrzad and Mcdermid, Christopher, “Availability and load balancing in cloud computing”, International Conference on Computer and Software Modeling, Singapore, chaczko2011availabilit

[6] R. Alonso-Calvo, J. Crespo, M. Garcia-Remesal, A. Anguita, V. Maojo, “On distributing load in cloud computing: A real application for very-large image datasets”, Procedia Computer Science 1 (1) (2010) pages 2669–2677.

[7] S.-C.Wang, K.-Q. Yan, S.-S.Wang, C.-W. Chen, “A three-phases scheduling in a hierarchical cloud computing network”, in: Communications and Mobile Computing (CMC), 2011 Third International Conference on, IEEE, 2011, pp. 114–117.

[8] Y. Hu, R. Blake, D. Emerson, “An optimal migration algorithm for dynamic load balancing”, Concurrency: Practice and Experience 10 (6)(1998) pages 467–483

[9] Etminani, Kobra and Naghibzadeh, M, “A min-min max-min selective algorihtm for grid task scheduling, Internet, 2007. ICI 2007. 3rd IEEE/IFIP International Conference in Central Asia on, 2007.

[10] Liu, Gang and Li, Jing and Xu, Jianchao, “ An Improved Min-Min Algorithm in Cloud Computing” page{47--52}, year{2013}, organization Springer

[11] U. Bhoi, P. N. Ramanuj, “ Enhanced max-min task scheduling algorithm in cloud computing”. pages={2319--4847}

[12] H. Chen, F. Wang, N. Helian, G. Akanmu, “User-priority guided min-min scheduling algorithm for load balancing in cloud computing”, pages={1--8}, year={2013}, organization={IEEE}

[13] Buyya, Rajkumar and Ranjan, Rajiv and Calheiros, Rodrigo N, “ modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities”, pages{1--11}, year{2009}, {IEEE}

[14] Zhu, Linan and Li, Qingshui and He, Lingna, “Study on Cloud Computing Resource Scheduling Strategy Based on the Ant Colony Optimization Algorithm., International Journal of Computer Science Issues (IJCSI), year 2012, volume 9

[15] Gao, Yongqiang and Guan, Haibing and Qi, Zhengwei and Hou, Yang and Liu, Liang, “A multi-objective ant colony system algorithm for virtual machine placement in cloud computing, Journal of Computer and System Sciences, year 2013, vol79, ppages1230—1242

[16] Mishra, Ratan and Jaiswal, Anant, “Ant colony Optimization: A Solution of Load balancing in Cloud., International Journal of Web \& Semantic Technology, year 2012, vol 3

[17] Li, Kun and Xu, Gaochao and Zhao, Guangyu and Dong, Yushuang and Wang, Dan, “Cloud task scheduling based on load balancing ant colony optimization, Chinagrid Conference (ChinaGrid), 2011 Sixth Annual, year 2011

[18] O. Elzeki, M. Reshad, M. Elsoud, “Improved max-min algorithm in cloud computing, International Journal of Computer Applications”vol 5(12)(2012)pages22–27

BER Analysis of MIMO-OFDM System Using Various Modulation Schemes

Vishal Sharma[1] and Gurpreet Singh[2]
[1,2]Department of Electronics & Comm. Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, India
E-mail: [1]er_vishusharma@yahoo.com, [2]er_gsandhu@yahoo.co.in

***Abstract*—Multiple transmit and receive can be used to form multiple input multiple-output (MIMO) channels to increase the capacity and data rate. The key advantage of employing multiple antennas is to get reliable performance through diversity and the achievable higher data rate through spatial multiplexing. In MIMO system, information can be transmitted and received from multiple antennas simultaneously since the fading for each link between a pair of transmit and receive antennas can usually be considered to be sovereign, so the probability that the information is detected precisely is higher. Fading of the signal can be diminished by different diversity techniques, where the signal is transmitted through multiple independent fading paths in terms of the time, frequency or space and combined constructively at the receiver. In this paper, Bit Error Rate (BER) of MIMO-OFDM system using various modulation schemes has been analyzed .**

Keywords: BER, MIMO, OFDM

I. Introduction

Multiple-Input Multiple-Output (MIMO) systems have changed wireless communications technology with the potential gains in capacity when multiple antennas at both transmitter and receiver ends of a communications systems are used. Fresh techniques were essential to grasp these gains in existing and new systems which account for the extra spatial dimension. MIMO technology has been adopted in multiple wireless standards, including Wi-Fi, WiMAX and suggested for future systems. Orthogonal frequency division multiplexing (OFDM) is an efficient technique to diminish ISI. OFDM [1] is a frequency division multiplexing (FDM) scheme utilized as a digital multicarrier modulation method [2]. In other words OFDM is frequency division multiplexing of multi-carriers those are orthogonal to each other means they are placed precisely at the nulls in the modulation spectra of each other. This makes OFDM [3] spectrally more efficient. In OFDM data is divided into a number of parallel data streams or sub-channels, one for every sub carriers which are orthogonal to each other though they overlap spectrally. Every sub-carrier is modulated with a conventional modulation scheme at a low symbol rate maintaining entire data rates similar to conventional single-carrier modulation schemes in the same bandwidth.

In today's development MIMO is very practical and beneficial with the combination of OFDM system. Utilizing the flexibility of MIMO systems in order to have high data rates is an especially smart research topic for future development scheme designs and their applications. Multiple-input multiple-output (MIMO) systems tender much larger channel capacity over traditional single-input single-output (SISO) systems. Lately, many transmit beam forming algorithms have been developed to utilize the high capacity in the MIMO systems [4][5]. Additionally, in MIMO systems, after selecting the group of users with the currently utmost possible rates are determined by a packet scheduler in each time-slot, they are needed to be assigned to the transmitter's antennas in a way that the maximum throughput in the system can be achieved. Diversity techniques like space-time coding have received attention due to their capability to offer higher spectral efficiency than conventional systems [6]. When these techniques are applied in a frequency-selective channel, a space-time equalizer is needed at the receiver to reimburse for the inter symbol interference (ISI). MIMO systems now comprise of a chief research area in telecommunications. It is also being considered as the one of the technologies that have a possibility to resolve the restricted access of traffic capacity in the forthcoming broadband wireless Internet access networks such as Universal Mobile Telephone Services(UMTS) and beyond Antenna selection has been recommended for enhanced performance in correlated fading [7][8]. Assuming that the number of RF chains is lesser than the number of antennas, the antenna selection algorithms decide the most favorable subset of transmit and receive antennas based on minimum error rate. The dimension of the active subsets of transmit and receive antennas is fixed by the number of RF chains, Per-antenna rate control which is equivalent to the power allocation, is applied to uncorrelated fading channels in [9], there it is shown that per-antenna rate control at the fading rate almost achieves capacity. Though, adaptation at the fading rate may be tricky to achieve in practice owing to inaccuracies in channel estimates and feedback delays. The remaining of this paper is organized as follows. MIMO and OFDM is discussed in Part II. In Part III model description and observations is done .In Part IV brief conclusion is given. Finally references are given.

II. MIMO System and OFDM

Conventionally, multiple antennas (at one side of the wireless link) have been used to carry out interference cancellation and to realize diversity and array gain in the course of coherent combining. The use of multiple antennas at both sides of the link bids an additional fundamental gain, known as spatial multiplexing gain, which results in raised spectral efficiency. Spatial multiplexing yields a linear capacity boost as compared to systems with a single antenna at one or both sides of the link, at no additional power or bandwidth spending [10–11]. The corresponding gain is accessible if the propagation channel exhibits affluent scattering and can be realized by the synchronized transmission of independent data streams in the same frequency band. The receiver exploits dissimilarities in the spatial signatures brought by the MIMO channel onto the multiplexed data streams to separate the different signals, thus realizing a capacity gain.

Diversity results in improved link reliability by rendering the channel "less fading" and by escalating the sturdiness to co-channel interference. Diversity gain is achieved by transmitting the data signal over multiple independently fading dimensions in time, space, and frequency and by performing appropriate combining in the receiver. Spatial diversity is mainly attractive as compared to time or frequency diversity, because it does not tempt an expenditure in transmission time or bandwidth, respectively. Space-time coding [12] understands spatial diversity gain in systems with multiple transmit antennas with no requirement of channel knowledge at the transmitter. Multiple antennas can be used at one or both sides of the wireless link to cancel or lessen co-channel interference, and therefore improve cellular system capacity. MIMO technology will largely be used in broadband systems that show frequency-selective fading and, therefore, intersymbol interference (ISI). OFDM modulation turns the frequency-selective channel into a set of parallel flat fading channels and is an attractive way of coping with ISI. The basic principle that lie behind OFDM is the introduction of a guard interval, known as cyclic prefix (CP), which is basically a copy of the last part of the OFDM symbol, and need to be long enough to house the delay spread of the channel [13]. The employment of the CP converts the action of the channel on the transmitted signal from a linear convolution into a cyclic convolution, so that the resulting overall transfer function can be diagonalised by the use of an IFFT at the transmitter and an FFT at the receiver end. As a result, the overall frequency-selective channel is converted into a set of parallel flat fading channels, which in turn radically simplifies the equalization job. On the other hand, as the CP carries redundant information, it invites a loss in spectral efficiency, which is usually kept at a maximum of 25 percent. OFDM has tighter synchronization requirements than single-carrier (SC) modulation and direct-sequence spread spectrum (DSSS), is more vulnerable to phase noise, and suffers from a larger peak-to-average power ratio. Multiple access and broadcasting is basically different in systems with multi-antenna terminals and base stations compared to systems with single-antenna terminals base stations or both. The fundamental reason is that realizing spatial- multiplexing gain needs the users to bump in signal space. This favors collision-based multiple-access schemes like code division multiple-access (CDMA) over orthogonal multiple-access schemes such as frequency division multiple access (FDMA) or time division multiple-access (TDMA).

The bit error rate or bit error ratio (BER) is the number of bit errors divided by the total number of transferred bits during a calculated time interval. BER is a unit less performance measure, frequently expressed in the form of percentage. The bit error probability is the expectation value of the BER. The BER can be considered as an estimated guess of the bit error probability. This estimate is accurate for a long time interval and a high number of bit errors. Calculating the bit error rate assists to choose the appropriate forward error correction codes. Given that most such codes only bit-flips, but not bit – insertions or bit detection, the hamming distance metric is the suitable way to compute the number of bit errors. The BER may be enhanced by choosing strong signal strength by choosing a slow and robust modulation scheme or line coding scheme, and also by applying channel coding schemes such as redundant forward error correction codes.

III. Model Description & Observations

For simulation of MIMO-OFDM, the system was modeled using Orthogonal Space Time Block Coding (OSTBC) technique which proposes spatial diversity gain to attain higher data rates. The OSTBC technique is an smart scheme as it can results in the full spatial diversity order and have symbol-wise maximum-likelihood (ML) decoding. The employ of Orthogonal Space-Time Block Codes (OSTBC) to achieve diversity gains in a multiple-input multiple-output (MIMO) communication system using OFDM is established. The simulation creates a random binary signal, modulates it by means of a phase shift keying (PSK) technique, and then encodes the waveform using a rate orthogonal space-time block code for transmission over the fading channel. The fading channel models six independent links, owing to the three transmit by two receive antennae configuration as single-path Rayleigh fading processes. The simulation attaches white Gaussian noise at the receiver. Then, it combines the signals from both receive antennas into a single stream for demodulation. For this combining process, the model assumes perfect knowledge of the channel gains at the receiver. For an IFFT transformation, a cyclic prefix (CP) is added to the signals and gets transmitted through the channel. At the receiver, the CP is removed if present and afterward the FFT is applied to reconstruct the signal. The FFT length considered for these simulations is 64. Total 128 subcarriers have been taken. OFDM symbol rate is 7.5 Ksps and the symbol period is 10^{-3}s. The system is designed for transmission of data over three transmit antennas and two receive antennas (3 x 2) using independent Rayleigh fading per link. Simulations were done using MATLAB to evaluate the performance of the system. The simulation compares the demodulated data with the original transmitted data, computing the bit error rate. The results have been compared for the BER performance as a function of signal to noise ratio (Eb/N0) using BPSK, its higher order and QAM modulation. In case of BPSK modulation, to have BER = 10^{-5}, MIMO OFDM requires 15 dB of Eb/No, while theoretical value of SNR required is 13 dB for achieving BER of 10^{-5}. On dealing with QPSK, high SNR is required to achieve BER of even 10^{-4}. Also, it has been observed that 8-PSK performs very badly and achieved only 10^{-3} at 20 dB. Alternatively, with QAM for both the simulated systems, the SNR requirement is almost same to the BPSK scheme. It is clear from all the simulated modulation schemes as shown in Figures 1-4 that BPSK has better BER performance as compared to its higher orders with low power penalty over Rayleigh Fading channel. The following graphs illustrate the performance of MIMO based OFDM as compared to their theoretical counterparts. For the theoretical results, the E_bN_o is scaled by the diversity order (four in this case).

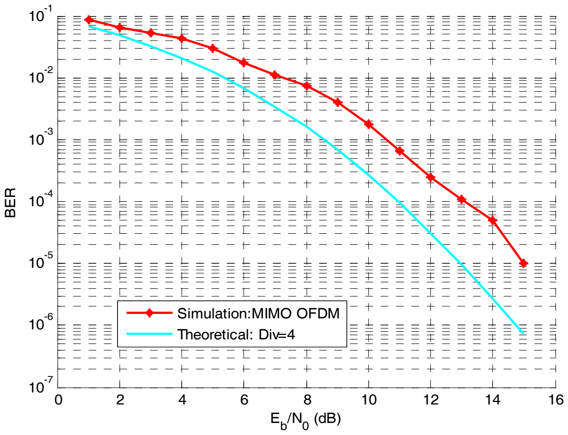


Fig. 1 Measured BER for OSTBC over 3 x 2 Rayleigh Fading Channel using BPSK

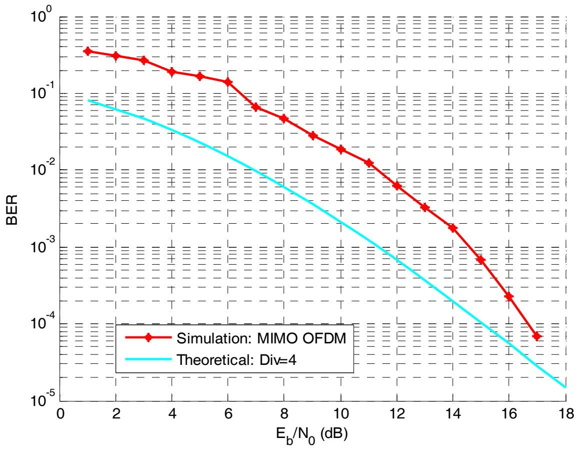


Fig. 2 Measured BER for OSTBC over 3 x 2 Rayleigh Fading Channel using QPSK

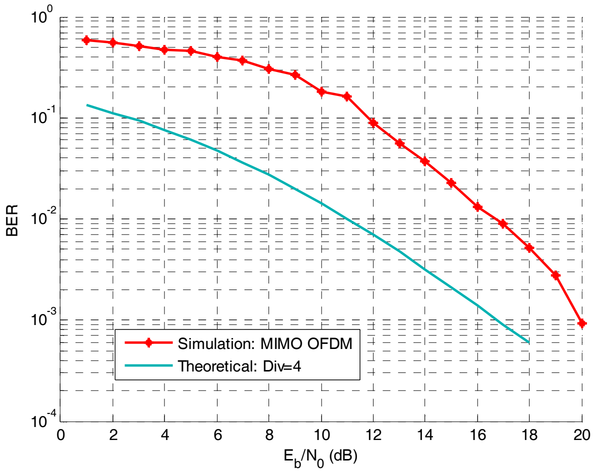


Fig. 3 Measured BER for OSTBC over 3 x 2 Rayleigh Fading Channel using 8-PSK

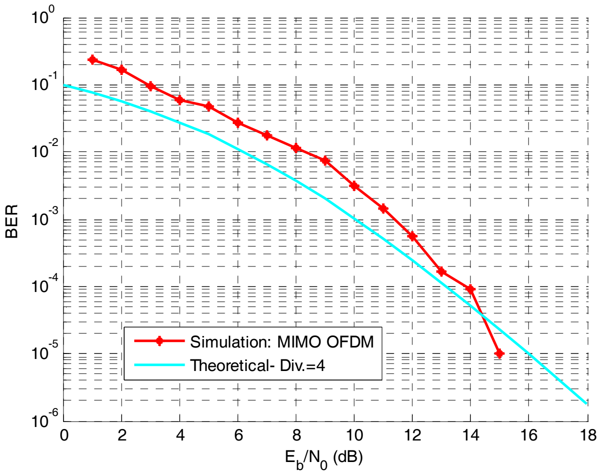


Fig. 4 Measured BER for OSTBC over 3 x 2 Rayleigh Fading Channel using QAM

IV. CONCLUSION

This work presents the performance of MIMO - OFDM system using various modulations techniques over Rayleigh fading channel. For higher orders of PSK schemes more SNR requirement is reported to target an acceptable BER over the simulated channel. Further, QAM requests fewer SNR as compared to PSK to achieve a suitable BER and is approximately same as that of BPSK. Lower order modulation schemes can be considered, but this is at the disbursement of data throughput. It is essential to balance all the available factors to realize a satisfactory bit error rate. Generally all the requirements are not achievable and some trade-offs are necessary.

REFERENCES

[1] F.B. Frederiksen and R.prasad, "An overview of OFDM and related techniques towards development of future wireless multimedia communications, " in proc. IEEE National conference on Radio and Wireless Communication conf. pp.19-22, Aug.2002.

[2] A.johan and CBingham, "Multicarrier modulation for datatransmission: An idea whose time has come, " IEEE commun.Mag., vol.28, no.5, pp.5-14, May 1990.

[3] Zhengdao Wang, "OFDM or single carrier block transmission, "IEEE Trans.on Comm., vol.12, no.3, pp.380-394, Mar 2004.

[4] P.Xia and G.Giannakis, "Design and analysis of transmit beamforming Based on limited-rate feedback, " IEEE Trans. Signal Processing.

[5] S.Zhou, Z.Wang, and G.Giannakis, "Quantifying the power loss when transmit beam forming relies on finite rate feedback, " IEEE Trans.

[6] V.Tarokh, N.Seshadri, and A. Calderbank, "Space-time Codes for High Data Rate Wireless communications Performance Criterion and Code Construction, "IEEETrains.Inform. Theory, vol.44, pp. 744-765, Mar.1998.

[7] D.Gore, R.Heath, and A.Paulraj, "Statistical antenna selection for spatial multiplexing systems, " in Proc.IEEE Int.Contr.Conf., New York, apr.2002.

[8] D. Gore and A.Paulraj, "MIMO antenna subset selection with space-time coding, " IEEE Trans.Signal Processing, vol.50, pp.2580-2588, Oct.2002.

[9] S.T. Chung, A.Lozano, and H.C.Huang, "Approaching eigenmode BLAST channel Capacity Using V-BLAST with rate and power feedback, " in Proc. Veh. Technol.Conf.Atlantic City, NJ, Oct. 2011, pp.915-919.

[10] A. J. Paulraj and T. Kailath, “Increasing Capacity in Wireless Broadcast Systems Using Distributed Transmission/ Directional Reception, ” U.S. Patent no. 5, 345, 599, 1994.

[11] I. E. Telatar, “Capacity of Multi-Antenna Gaussian Channels, ” *Euro. Trans. Telecommun.*, vol. 10, no. 6, Nov./Dec. 1999, pp. 585–95.

[12] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space- Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction, ” *IEEE Trans. Info. Theory*, vol. 44, no. 2, Mar. 1998, pp. 744–65.

[13] Mahesh Kumar Gupta and S. Tiwari, “Performance evaluation of conventional and wavelet based OFDM system, ” Elsevier, 2012.

Study of LEACH Routing Protocol for Wireless Sensor Networks

Reenkamal Kaur Gill[1], Priya Chawla[2] and Monika Sachdeva[3]
[1,2,3]Department of Computer Science and Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]reenkamalgill@gmail.com, [2]piyachawla12@gmail.com,
[3]monika.sal@rediffmail.com

***Abstract*—Wireless Sensor Network are in great demand from the recent years, as nowadays we have seen a wide growth of wireless devices including cellular phones, laptops, mobiles, PDA's etc. Wireless Sensor Networks consists of thousands of tiny sensor nodes. In a wireless sensor network a node is no longer useful when its battery dies, so to avoid this problem many protocols were introduced, but most of the rank is given to hierarchical routing protocols. In this paper, we analyze LEACH protocol, its phases, advantages and disadvantages and also various kinds of attacks on this routing protocol.**

Keywords: Wireless Sensor Networks, LEACH Protocol, Cluster, Cluster Head, Attacks

I. Introduction

A wireless sensor networks consist of tiny sensor nodes to monitor physical or environmental conditions such as temperature, pressure, sound, humidity etc. The network must possess self configuration capabilities as the positions of the individual sensor nodes are not pre-determined.

Routing strategies and security issues are a great research challenge now days in WSN but in this paper we will emphasize on the routing protocol. A number of routing protocols have been proposed for WSN but the most well known are hierarchical protocols like LEACH [1] and PEGASIS [2].

Hierarchical protocols are defined to reduce energy consumption by aggregating data and to reduce the transmissions to the Base Station. LEACH is considered as the most popular routing protocol that use cluster based routing in order to minimize energy consumption.

In this paper firstly we analyze LEACH protocol and then in the third section we will discuss the phases of LEACH protocol. In the fourth section we define various possible attacks on it and in the fifth section there are the advantages and disadvantages of LEACH. In the last section we compare LEACH with other protocols.

II. LEACH

Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is a TDMA based MAC protocol. The principal aim of this protocol is to improve the lifespan of wireless sensor networks by lowering the energy consumption required to create and maintain Cluster Heads. The operation of LEACH protocol consists of several rounds with two phases in each [3] [4]: Set-up Phase and Steady Phase. In the Set-up phase the main goal is to make cluster and select the cluster head for each of the cluster by choosing the sensor node with maximum energy. Steady Phase which is comparatively longer in duration than the set-up deals mainly with the aggregation of data at the cluster heads and transmission of aggregated data to the Base station.

III. Phases of LEACH

As described earlier the operation of LEACH consists of several rounds with two phases in each round. Working of LEACH starts with the formation of clusters based on the received signal strength.

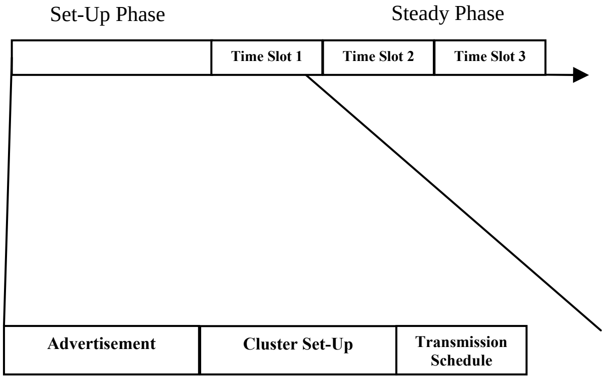


Fig. 1 Time Line Operation of LEACH

The algorithm for LEACH protocol is as follows:

The first phase of LEACH is Set-up phase and it has three fundamental steps.

1. Cluster Head advertisement
2. Cluster setup
3. Creation of Transmission Schedule

During the first step cluster head sends the advertisement packet to inform the cluster nodes that they have become a cluster head on the basis of the following formula [5]:

Let x be any random number between 0 and 1.

Where n is the given node, p is the probability, r is the current round, G is the set of nodes that were not cluster heads in the previous round, T(n) is the Threshold.

$$T(n) = \begin{cases} \dfrac{P}{1 - P[r * \mathrm{mod}(1/P)]} & if \;\; n \in G \\ 0 & otherwise, \end{cases}$$

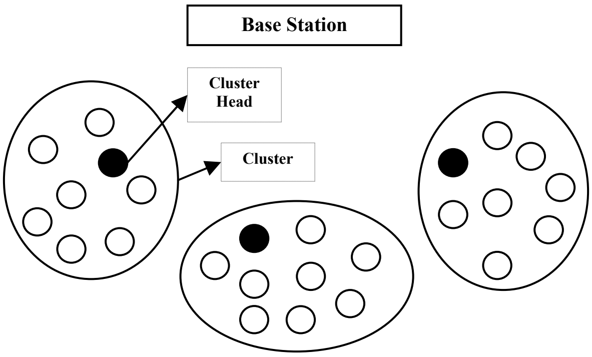


Fig. 2 Cluster Formation in LEACH.

The node becomes cluster head for the current round if the number is less than threshold T(n). Once the node is elected as a cluster head it cannot become cluster head again until all the nodes of the cluster have become cluster head once. This helps in balancing the energy consumption.

In the second step, the non cluster head nodes receive the cluster head advertisement and then send join request to the cluster head informing that they are the members of the cluster under that cluster head as shown in Fig. 2 [6].

These non cluster head nodes saves a lot of energy by turning off their transmitter all the time and turn it ON only when they have something to transmit to the cluster head.

In the third step, each of the chosen cluster head creates a transmission schedule for the member nodes of their cluster. TDMA schedule is created according to the number of nodes in the cluster. Each node then transmits its data in the allocated time schedule.

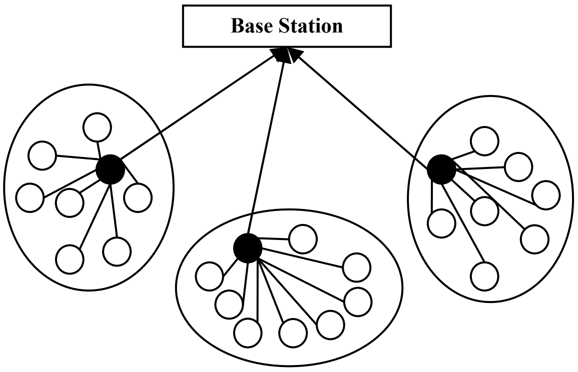


Fig. 3 Steady Phase in LEACH.

The second phase of LEACH is the Steady phase during which the cluster nodes send their data to the cluster head. The member sensors in each cluster communicate only with the cluster head via a single hop transmission. The cluster head then aggregates all the collected data and forwards this data to the base station either directly or via other cluster head along with the static route defined in the source code as shown in Fig. 3[7]. After the certain predefined time, which is decided beforehand, the network again goes back to the Set-up phase.

IV. Attacks on LEACH

LEACH protocol is threatened by the following types of attacks which degrade the performance of LEACH by dropping, altering, spoofing or replying the packets.

A. *Sybil Attack*

Most of the peer to peer networks face security threats due to Sybil attack [8], [9]. This attack is the most difficult attack to detect. In this attack, malicious node uses the identity of many other legitimate nodes to gain the data exchanged between the legitimate nodes. It affects the network by dropping vital packets, increasing traffic, lowering network lifetime etc. Encryption and authentication techniques can be used to prevent wireless sensor network from the Sybil attack.

B. *Selective Forwarding*

LEACH protocol is also susceptible to selective forwarding attack. In this kind of attack a malicious node places itself in the path where data is exchanged between the two legitimate nodes. It collects the data and instead of forwarding this node drops all the data. It is the case where the malicious node can easily be detected. The worst scenario of this attack is that when malicious node does not discard the entire data, but selectively forwards some of the non vital information. In this case it is very difficult detect the malicious node.

C. *HELLO Flooding Attack*

In many protocols sometimes it is required for node to transmit HELLO packets to advertise itself to its neighboring nodes. The nodes receiving these packets assume that it is within the range of the sender. But in case of malicious node, it continuously keeps on sending the HELLO packets and thus increases the network traffic and causes collisions. It also consumes the energy of the sensor nodes when these nodes receive large amount of HELLO packets continuously and thus lowering the lifetime of the wireless sensor networks. This type of attack is known as HELLO Flood attack [10].

V. Advantages and Disadvantages of LEACH

The various advantages [11] of LEACH protocol are:

1. The Cluster Heads aggregates the whole data which lead to reduce the traffic in the entire network.
2. As there is a single hop routing from nodes to cluster head it results in saving energy.
3. It increases the lifetime of the sensor network.

4. In this, location information of the nodes to create the cluster is not required.
5. LEACH is completely distributed as it does not need any control information from the base station as well as no global knowledge of the network is required.

Besides the advantages of LEACH it also has some demerits [11], [12] which are as follows:

1. LEACH does not give any idea about the number of cluster heads in the network.
2. One of the biggest disadvantage of LEACH is that when due to any reason Cluster head dies, the cluster will become useless because the data gathered by the cluster nodes would never reach its destination i.e. Base Station.
3. Clusters are divided randomly, which results in uneven distribution of Clusters. For e.g. some clusters have more nodes and some have lesser nodes. Some cluster heads at the center of the cluster and some cluster heads may be in the edge of the cluster; this phenomenon can cause an increase in energy consumption and have great impact on the performance of the entire network.

VI. Conclusion

Wireless Sensor Networks would be of great use in future mission applications. If we analyze the previous research, we could observe that a lot of work is being carried out on routing i.e. what is the best optimal path for the nodes to communicate with each other. In this paper, we have also discussed LEACH routing protocol. Basically how does it works has been explained above with its advantages and disadvantages. LEACH protocol is also vulnerable to various kinds of attacks which have been described above.

References

[1] W. Heinzelman, A. Chandrakasan, H. Balakrishnan, "*Energy-efficient communication protocol for wireless sensor networks*", in: Proceeding of the Hawaii International Conference System Sciences, Hawaii, January 2000.

[2] S. Lindsey, C.S. Raghavendra, "*PEGASIS: power efficient gathering in sensor information systems*", in: Proceedings of the IEEE Aerospace Conference, Big Sky, Montana, March 2002.

[3] Rajesh Patel, Sunil Pariyani, Vijay Ukani," *Energy and hroughput Analysis of Hierarchical Routing Protocol(LEACH) for Wireless Sensor Networks"*, International Journal of Computer Applications Volume 20- No. 4 (April 2011).

[4] Yuh Ren Tsai, "*Coverage Preserving Routing Protocols for Randomly Distributed Wireless Sensor Networks"*, IEEE Transactions on Wireless Communications, Volume 6- No. 4 (April 2007).

[5] Amrinder Kaur, Sunil Saini," *Simulation of Low Energy Adaptive Clustering Hierarchy Protocol for Wireless Sensor Network*," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue7, July 2013.

[6] Ian F. Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, Erdal Cayirci: *"A Survey on Sensor Networks"*, IEEE Communication Magazine, pp. 102-114(August 2002).

[7] Rajesh Patel, Sunil Pariyani, Vijay Ukani," *Energy and Throughput Analysis of Hierarchical Routing Protocol(LEACH) for Wireless Sensor Networks"*, International Journal of Computer Applications Volume 20- No. 4 (April 2011).

[8] Douceur, J. "The Sybil Attack", 1st International Workshop on Peer-to-Peer Systems (2002).

[9] Newsome, J., Shi, E., Song, D, and Perrig, A, "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.

[10] Karlof, C. and Wagner, D., "Secure routing in wireless sensor networks: Attacks and countermeasures", Elsevier's Ad Hoc Network Journal, Special Issue on Sensor Network Applications and Protocols, September 2003, pp. 293-315.

[11] Parul Kansal, Deepali KAnsal, Arun Balodi,*"Compression of Various Routing Protocol in Wireless Sensor Networks"*, International Journal of Computer Applications Volume 5-No. 11(August 2010) .

[12] M. Bani Yassein, A. AL-zou'bi, Y. Khamayseh, W. Mardini, *"Improvement on LEACH Protocol of Wireless Sensor Networks"*, International Journal of Digital Content Technology and its Applications Volume 3- No. 2 (June 2009).

Current Methods in Medical Image Segmentation: A Review

Ramandeep Kaur[1] and Jaswinder Kaur[2]
[1,2]Department of Electronics & Comm. Engineering
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab
E-mail: [1]ramancheema92@gmail.com, [2]jaswinder.ece@gmail.com

***Abstract*—In this paper, different approaches are discussed for medical image segmentation. These are based on thresholding, learning, modeling and automatic fuzzy method. Segmentation techniques, discussed under these approaches are used in different applications. In identification of brain lesions, vessel lumen segmentation and histopathology cancer image segmentation. Further used in tissue segmentation based upon image processing chain optimization, combining graph cut and oriented Active Appearance Model (AAM) and in brain image segmentation by using fuzzy symmetry. These techniques overcome various limitations of conventional medical image segmentation techniques.**

Keywords: CTA, MRA, MRI, Segmentation, Thresholding

I. Introduction

Image segmentation is the process of partitioning a digital image into multiple segments or sets of pixels, which are also known as super pixels. Basically segmentation is used to simplify and/or analyze images [1] [2]. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. In image segmentation process a label is assigned to every pixel in an image and pixels with the same label share certain characteristics. Set of segments obtained as a result of image segmentation and these segments collectively cover the entire image. Image segmentation using thresholding was not satisfactory in medical imaging. Due to the high dimensionality of the image relative to smaller sample sizes direct estimation of the statistical variation of the entire volumetric image was challenged, vascular segmentation was not easily possible and automated reconstruction of cortical surface was also the most challenging problem in the analysis of human brain Magnetic Resonance Imaging (MRI). Labeling a histopathology image as having cancerous regions or not was a critical task in cancer diagnosis. Only 2D models were used, 3D models were not compatible in medical imaging.

II. Thresholding Based Segmentation

Thresholding based segmentation, in which one threshold value is used to select the area of interest and this threshold value can be selected by using prior knowledge or from image information. Further threshold approach can be edge based, region based or hybrid. In edge based approach, edge information is required. Canny edge detector and Laplacian edge detectors work on this approach. Canny edge detector uses the threshold of gradient magnitude to find the potential edge pixels and suppresses them through the procedures of the non - maximal suppression and hysteresis thresholding. In this, the detected edges are consisted of discrete pixels, may be incomplete or discontinuous. So, it is necessary to apply post-processing like morphological operation to connect the breaks or eliminate the holes. Pixels inside a structure tend to have similar intensities and from this observation idea of region-based algorithms developed. Region growing algorithm is a typical algorithm of this type. In this algorithm firstly initial seeds are selected then it search for the neighbored pixels whose intensities are inside the intervals defined by the thresholds and then merge them to expand the regions. Statistical information and a priori knowledge can be incorporated to the algorithms to eliminate the dependence on initial seeds and make the algorithm automatic. For example, a homogeneity criterion was introduced in [3], which made the region growing algorithms adaptive for the different locations of initial seeds.

These algorithms mainly rely on the image intensity information, so they are hard to handle the partial volume effects and control the leakage. Watershed algorithms are typical example of this algorithm [4], which combines the image intensity with the gradient information. In this algorithm, gray scale images are considered as reliefs and the gradient magnitude of each pixel is treated as elevation. Watershed lines are defined to as the pixels with local maximum of gradient magnitude. Segmentation procedure is used to construct watersheds during the successive flooding of the gray value relief. Watershed algorithms can achieve better results due to the combination of image information, but when the images are noisy or the objects themselves have low signal-to-noise ratio these algorithms tend to over-segmentation. Hybrid threshold-based algorithms can further combine with other techniques to perform the segmentation [5]. Due to the noise influence and partial volume effect algorithms based on threshold are seldom used alone because the edges of organs or structures in medical images are usually not clearly defined.

III. Learning Based Segmentation

In learning based approach, there may be use of statistical learning, supervised, unsupervised, can be weakly supervised also. Techniques based upon this approach are following

A. *Individualized Statistical Learning from Medical Image Databases*

This method works on comparison of normative set of images with other images and statistical variation is estimated; as a result abnormalities are identified as deviations from normality. Direct estimation of the statistical variation of the entire image is not possible because of high-dimensionality of images relative to smaller sample sizes [6]. Similarly, large numbers of lower dimensional subspaces are iteratively sampled that capture image characteristics ranging from fine and localized to coarser and more global. Within each subspace, a "target-specific" feature selection strategy is applied to further reduce the dimensionality. Marginal probability Density Functions of selected features (by considering only imaging characteristics present in a test subject's images) are estimated through Principal Component Analysis (PCA) models, in conjunction with an “estimability” criterion that limits the dimensionality of estimated probability densities according to available sample size and underlying anatomy variation.

A test sample is iteratively projected to the subspaces of these marginal’s as determined by PCA models, and its trajectory tells about potential abnormalities. The method is used for segmentation of various brain lesion types, and for simulated data on which superiority of the iterative method over straight PCA is demonstrated.

With this method problem of high dimensionality of the image domain relative to the typically available sample sizes get solved by introducing an iterative method for sampling subspace, by incorporating many more images from healthy subjects into the model, the performance of this method could be improved.

B. *3D Vessel Lumen Segmentation Techniques: Features and Extraction Schemes*

The segmentation of vascular structures is particularly valuable for diagnosis assistance, treatment and surgery planning. Segmentation is fundamental step for the accurate visualization of vessels from complex datasets and for the quantification of pathologies. Unfortunately, most angiographic clinical routines still rely heavily on manual operations. Modern 3D imaging modalities are Computed Tomography Angiography (CTA) and Magnetic Resonance Angiography (MRA), manual segmentation can quickly add up to hours of processing. In this context, automatic and semi-automatic image processing tools aim at easing and speeding up reviewing tasks, reducing the amount of manual interaction and lowering inter-operator variability. Vascular segmentation is an especially specific and challenging problem [7].

Besides general, acquisition-dependent considerations about contrast, resolution, noise and artifacts, vascular networks can be particularly complex structures. Blood vessels potentially exhibit high variability of size and curvature. Their appearance and geometry can be perturbed by stents, calcifications, aneurysms, and stenos. Finally, they are often embedded in complex anatomical scenes, surrounded by other organs. Basic components of this method are appearance and geometric models, image features, extraction schemes.

Models correspond to the prior assumptions made on the target vessels, e.g., elongation and hyper-intensity. Features are the vessel dedicated image measures used to estimate the models on the image, e.g. local intensity curvatures. Finally, the extraction scheme represents the algorithmic core of a vascular segmentation method. A way to improve existing algorithms in terms of both performance and automation can thus be the design of new sequential combinations.

C. *Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces for the Brain*

LOGISMOS-B, based on probabilistic tissue classification, gradient vector flows and the LOGISMOS graph segmentation framework. Quantitative results on Magnetic Resonance Imaging (MRI) datasets from both healthy subjects and multiple sclerosis patients using a total of 16, 800 manually placed landmarks illustrate the excellent performance of the algorithm with respect to spatial accuracy. Even in the presence of multiple sclerosis lesions, average signed errors were only 0.084 mm and 0.008 mm for white and gray matter respectively. Observation from statistical comparison gives that LOGISMOS-B produces a significantly more accurate cortical reconstruction than Free Surfer [8], the current state-of-the-art approach.

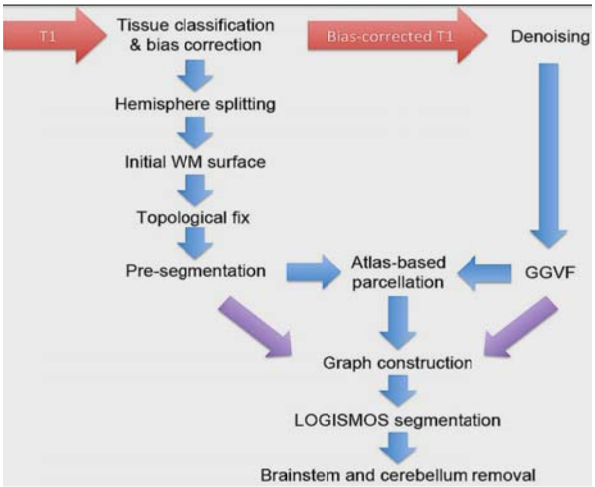


Fig. 1 Pipeline Overview

Furthermore, LOGISMOS-B enjoys a run time that is less than a third of that of Free Surfer, which is both substantial, considering the latter takes ten hours per subject, and this method provides statistically significant speedup. Cortical segmentation method LOGISMOS-B1 consists of four main steps that are Pre processing of raw image to create a rough preliminary segmentation; graph construction; LOGISMOS segmentation; and post-processing for removal of brain stem and cerebellum. An overview of the pipeline is given in Fig. 1. This method offers improved anatomic accuracy and dramatically reduced computational requirement.

D. *Weakly Supervised Histopathology Cancer Image Segmentation and Classification*

Labeling a histopathology image as having cancerous regions or not is a critical task in cancer diagnosis. It is also clinically important to segment the cancer tissues and cluster them into various classes. Existing supervised approaches for image classification and segmentation require detailed manual annotations for the cancer pixels, which are time-consuming to obtain.

A new learning method, Multiple Clustered Instance Learning (MCIL) for histopathology image segmentation. MCIL method simultaneously performs image-level classification (cancer vs. non-cancer image) as shown in Fig. 2. medical image segmentation (cancer vs. non-cancer tissue), and patch-level clustering (different classes [9].

Under this concept, Multiple Instance Learning (MIL) performs the above three tasks in an integrated framework. In addition, under this contextual constraints are introduced as a prior in MCIL, which further reduces the ambiguity in MIL. Experimental results on histopathology colon cancer images and cytology images demonstrate the great advantage of MCIL over the competing methods.

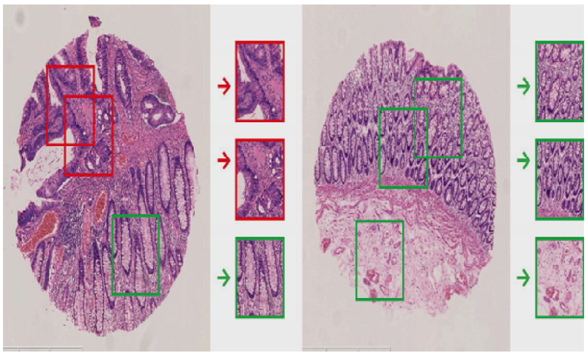


Fig. 2 Cancer and Non Cancer Images

The advantages of MCIL are evident over the state-of the- art methods that perform the individual tasks, which include easing the burden of manual annotation in which only image-level label is required and perform image-level classification, pixel-level segmentation and patch-level clustering simultaneously. In addition, we introduce contextual constraints as a prior for MCIL which reduces the ambiguity in MIL. MCIL and contextual MCIL are able to achieve comparable results in segmentation with an approach of full pixel-level supervision in experiment. This will inspire future research in applying different families of joint instance models to the framework of MIL/MCIL, as the independence assumption might be loose.

IV. Model Based Segmentation

A. *Statistical Shape Models for 3D Medical Image Segmentation*

Nowadays, model-based segmentation approaches have been established as one of the most successful methods for image analysis by matching a model. Model contains information about the expected shape and appearance of the structure of interest to images, in this segmentation is conducted in a top-down fashion. This method is more stable against local image artifacts and perturbations than conventional low-level algorithms. While a single template shape is an adequate model for industrial applications where mass produced, rigid objects need to be detected, this method is prone to fail in case of biological objects due to their considerable natural variability [10]. Information about common variations thus has to be included in the model.

A straight-forward approach to gather this information is to examine a number of training shapes by statistical means, leading to Statistical Shape Models (SSMs). Training data for SSMs in the medical field will most likely consist of segmented volumetric images. Depending on the segmentation method used, the initial representation can be in form of binary, voxel data, fuzzy voxel data (e.g. from probabilistic methods), or surface meshes. Data originating from other sources of acquisition, e.g. surface scanning might be represented differently. In any case, all shape representations can be converted into each other, and the choice of shape representation is the first fundamental decision when designing statistical shape models. Most of the subsequent steps depend on this initial decision, and many methods are technically limited to certain representations, include constructing shape models which is basically extracting of the mean shape and a number of modes of variation from a collection of training samples. Obviously, the methods employed strongly depend on the chosen shape representation. After this Shape correspondence that is basically modeling the statistics of a class of shapes requires a set of appropriate training shapes with well-defined correspondences.

Depending on the chosen representation, the methods of how to best define these correspondences vary. In any case, establishing dense point correspondences between all shapes of the training set is generally the most challenging part of 3D model construction, and at the same time one of the major factors influencing model quality (the other one being the local gray-value appearances).After construction the model is fitted to new, previously unseen data. For this purpose, a model of the appearance of the structure of interest is required to be trained from sample data, due to the large size of the search space in 3D, most methods applied to locate an SSM in new image data use local search algorithms that require an initial estimate of the model pose.

B. *Medical Image Segmentation by Combining Graph Cuts and Oriented Active Appearance Models*

This method is combination of active appearance model (AAM), Live Wire (LW) and Graph Cuts (GCs) for abdominal 3D segmentation of organs. This method consists of two phases training and segmentation. In training phase AAM algorithm is constructed and LW boundary cost functions and GC parameters are estimated, and segmentation phase consists two main parts recognition or initialization and delineation [11].

In the recognition step, a pseudo-3-D initialization strategy is employed in which the pose of the organs is estimated slice by slice via a multi object OAAM (MOAAM) method. In the delineation part object shape information generated from the initialization step is integrated into GC cost computation.

C. *Fuzzy C-Mean (FCM) Method for Segmentation of Brain MRI Image*

In this method, with the help of Self Organizing Map(SOM) clustering algorithm initial cluster centers are selected, after many iterations of this algorithm final cluster centre is obtained. The winning neural units and their corresponding weight vectors from each layer result in an abstraction tree. A particular region of the image at a certain level of abstraction is represented with one node of this abstraction tree [12]. Under this segmentation is performed on demand by transverseing the abstraction tree in Breadth-First Search(BFS) manner starting from root node until certain criteria is satisfied. If the sum of variances of weight vector divided by size of weight vectors is less than element of weight vector if the size of abstraction tree is expanded else the node is labeled as closed node and regions corresponds to closed nodes constitute a segmented image.

D. *LVQ Method for Segmentation of Brain MRI Image*

Linear Vector Quantization (LVQ) technique is supervised learning technique obtain decision boundaries based upon training data .In this method three layers are there input, competitive and output layer [12]. Input data is classified in the competitive layer and then those classes or patterns are mapped to the target class in the output layer, under this winner neuron is selected based upon the Euclidean distance then weights of this winner neurons can be adjusted by using different algorithms.

E. *SOM and Hybrid SOM Method for Segmentation of Brain MRI Image*

In SOM method, firstly find the winning neuron and secondly updating weight of the neuron and its neighboring pixels based upon input [12]. Hybrid SOM combines self organization and topographic mapping technique.

F. *Markov Random Field (MRF) Model and Fast Fourier Transform (FFT) Based Segmentation for Segmentation of Brain MRI Image*

In Markov Random Field model neighborhood information is used, because most neighborhood pixels are in same class as a result influence of noise decreased [12]. FFT based segmentation used in brain segmentation because in all tumors boundaries between active and necrotic part are not clear, for this radix 4 FFT partitions Discrete Fourier Transform (DFT) into four quarter length DFT's of groups of every fourth time sample, total computational cost reduced by these FFT outputs which are reused for computing the output.

G. *Tissue Segmentation in Medical Images Based on Image Processing Chain Optimization*

Differential evolution method is purposed to optimize an image processing chain .In this method training is based upon three sample images provided by an expert [13]. Mainly Differential Evolution(DE) method is population based optimization method, idea behind DE is generating trial parameter vectors, for every vector in the population, DE selects randomly two other vectors, subtract them and add the weighted difference to randomly chosen third vector(base vector) to produce mutant vector, cross over rate (user defined value) is used for every vector in mutant population to control the fraction of parameter values that are copied from the mutant and target vector to trial vector, if trial vector have equal or lower fitness value than that of its target vector, then it replaces the target vector in next generation, otherwise target retain its place for at least one more generation, steps are repeated for every vector in population to generate new population.

This method trying to overlap gold images by well known images generated by experts and images processed with this technique.

V. AUTOMATIC FUZZY APPROACH FOR SEGMENTATION

A. *MRI Brain Image Segmentation by Fuzzy Symmetry Based Genetic Clustering Technique*

Automatic segmentation technique of MRI of brain using new fuzzy point symmetry based genetic clustering technique is proposed, which is able to evolve the number of clusters present in the data set automatically [14]. In this assignment of points to clusters are based on point symmetry based distance rather than the Euclidean distance and because of this, proposed algorithm Fuzzy Variable string length Genetic Point Symmetry(Fuzzy VGAPS) enable to identify any type of cluster irrespective of its shape size convexity and this method automatically evolve the clusters.

VI. Conclusion

In this paper, various segmentation techniques applied for medical images are briefly explained. All the discussed techniques overcome various limitations occurred in medical image segmentation like direct estimation of the statistical variation of the entire volumetric image, vascular segmentation and in the analysis of human brain magnetic resonance imaging(MRI) automated reconstruction of cortical surface was the most challenging problem. Labeling a histopathology image as having cancerous regions or not was a critical task in cancer diagnosis. 3D models were not compatible in medical imaging. Performance of these medical image segmentation techniques can be improved in future by incorporating many more images from healthy subjects into the models, by the design of new sequential combinations of different methods, for more optimization many algorithm can be added in pre and post processing phases, combination of graph cut and OAAM can be improved by introducing parallelization and by incorporating the spatial information.

References

[1] L. Barghout and L. Lee, “Perceptual information processing system, ” 2003.

[2] G. Srinivasan and G. Shobha, “An overview of segmentation techniques for target detection in visual images, ” in Proceedings of the 9th WSEAS International Conference on International Conference on Automation and Information. World Scientific and Engineering Academy and Society (WSEAS), 2008, pp. 511–518

[3] R. Pohle and K. D. Toennies, “Segmentation of medical images using adaptive region growing, ” in Medical Imaging 2001. International Society for Optics and Photonics, 2001, pp. 1337–1346.

[4] L. Najman and R. Vaillant, “Topological and geometrical corners by watershed, ” in Computer Analysis of Images and Patterns. Springer, 1995, pp. 262–269.

[5] S. A. Mani, W. Guo, M.-J. Liao, E. N. Eaton, A. Ayyanan, A. Y. Zhou, M. Brooks, F. Reinhard, C. C. Zhang, M. Shipitsin et al., “The epithelial-mesenchymal transition generates cells with properties of stem cells, ” Cell, vol. 133, no. 4, pp. 704–715, 2008.

[6] G. Erus, E. I. Zacharaki, and C. Davatzikos, “Individualized statistical learning from medical image databases: Application to identification of brain lesions, ” Medical image analysis, vol. 18, no. 3, pp. 542–554, 2014.

[7] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea, “A review of 3d vessel lumen segmentation techniques: Models, features and extraction schemes, ” Medical image analysis, vol. 13, no. 6, pp. 819– 845, 2009.

[8] I. Oguz and M. Sonka, “Logismos-b: Layered optimal graph image segmentation of multiple objects and surfaces for the brain, ” IEEE transaction, vol. 33, p. 6, 2014.

[9] Y. Xu, J.-Y. Zhu, E. I. Chang, M. Lai, Z. Tu et al., “Weakly supervised histopathology cancer image segmentation and classification, ” Medical image analysis, vol. 18, no. 3, pp. 591–604, 2014.

[10] T. Heimann and H.-P. Meinzer, “Statistical shape models for 3D medical image segmentation: A review, ” Medical image analysis, vol. 13, no. 4, pp. 543–563, 2009.

[11] X. Chen, J. K. Udupa, U. Bagci, Y. Zhuge, and J. Yao, “Medical image segmentation by combining graph cuts and oriented active appearance models, ” Image Processing, IEEE Transactions on, vol. 21, no. 4, pp. 2035–2046, 2012.

[12] S. Bandhyopadhyay and T. U. Paul, “Segmentation of brain MRI image–a review, ” International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 3, pp. 409– 413, 2012.

[13] S. Rahnamayan and Z. Mohamad, “Tissue segmentation in medical images based on image processing chain optimization, ” in International Workshop on Real Time Measurement, Instrumentation and Control, Toronto, 2010, pp. 1–9.

[14] S. Saha and S. Bandyopadhyay, “MRI brain image segmentation by fuzzy symmetry based genetic clustering technique, ” in Evolutionary Computation, 2007. CEC 2007. IEEE Congress on. IEEE, 2007, pp. 4417–4424.

Biogeography Based Optimization for Gain Maximization of Fifteen-element Yagi-Uda Antenna

Gagan Sachdeva[1], Dilpal Singh[2] and Satvir Singh[3]
[1]Rayat Bahra Group of Institutes, Mohali Campus, Punjab, India
[2]UIET, Panjab University, Chandigarh, India
[3]Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]gagan.sachdeva04@gmail.com, [2]dilpal.singh01@gmail.com, [3]drsatvir.in@gmail.com

***Abstract*—Biogeography-Based Optimization (BBO) is a recently introduced optimization technique based on science of biogeography, i.e., study of distribution of biological species over space and time. In BBO, potential solutions of a problem are grouped in integer vectors known as habitats. BBO uses migration operator for feature sharing among habitats and mutation operator to explore new features. Yagi-Uda antenna is a widely used directional antenna design due to various useful properties of high gain, low cost and ease of construction. Designing a Yagi-Uda antenna includes determination of element lengths and spacings between them to get desired radiation characteristics. The gain of Yagi-Uda antenna is hard to optimize as there is no analytical formula to determine gain directly, it makes relationship between antenna parameters and its characteristics highly complex and non-linear. In this paper, 15-element Yagi-Uda antenna is optimized for gain maximization using BBO. The results obtained by BBO are compared with Bi-Swarm optimization, Ellipsoid Algorithm and Genetic Algorithm (GA). BBO shows better results than other compared optimization techniques.**

Keywords: Biogeography Based Optimization, Yagi-Uda Antenna, Antenna Gain, Genetic Algorithm, Bi-Swarm Optimization, Ellipsoid Algorithm

I. INTRODUCTION

Antenna is an electrical device which converts electric signal into free space radiations and vice-versa. The various radiation characteristics that affect the design of an antenna are gain, impedance, bandwidth, frequency of operation, Side Lobe Level (SLL) etc. Yagi-Uda antenna is a widely used directional antenna design due to various desirable features, i.e., high forward gain, low cost and ease of construction. It is basically a parasitic linear array of parallel dipoles, one of which is energized directly by transmission line while the others act as parasitic radiators whose currents are induced by mutual coupling.

Yagi-Uda antenna was invented in 1926 by H. Yagi and S. Uda at Tohoku University [1] in Japan, however, published in English in 1928 [2]. The main objective, in design of Yagi-Uda antenna, is to find an optimum structure that meet certain radiation criteria like gain, impedance, SLL and beamwidth. However, due to its parasitic elements, it is extremely difficult to obtain an optimum design of Yagi-Uda antenna. Since its inception, Yagi-Uda antenna has been optimized several times for gain, impedance, SLL and bandwidth using different optimization techniques based on traditional mathematical approaches [3], [4], [5], [6], [7], [8], [9] and Artificial Intelligence (AI) techniques [10], [11], [12], [13], [14], [15], [16]. In 1949, Fishenden and Wiblin [17] proposed an approximate design of Yagi aerials for maximum gain, however, the approach was based on approximations. In 1959, Ehrenspeck and Poehler proposed a manual approach to maximize the gain of the antenna by varying various lengths and spacings of its elements [18].

Later on, with the availability of high performance computing, it became possible to optimize antennas numerically. Bojsen *et al.* in [4] proposed an optimization technique to find the maximum gain of Yagi-Uda antenna arrays with equal and unequal spacings between adjacent elements. Cheng *et al.*, in [7] and [8] have used optimum spacings and lengths to optimize the gain of a Yagi-Uda antenna. In [9], Cheng has proposed optimum design of Yagi-Uda antenna where antenna gain function is highly non-linear. The performance of these gradient based techniques depends on choice of initial solution.

In 1975, John Holland introduced Genetic Algorithms (GAs) as a stochastic, swarm based AI technique, inspired from natural evolution of species, to optimize arbitrary systems for certain cost function. Since then many researchers have used GAs to optimize Yagi-Uda antenna designs for gain, impedance and bandwidth separately [19], [10], [20] and collectively [11], [21], [22]. Jones *et al.*, in [10] have used GA to optimize Yagi-Uda antenna for various radiation characteristics and compared the result with steepest gradient method. Baskar *et al.* in [13], have used Comprehensive Learning Particle Swarm Optimization (CLPSO) to optimize Yagi-Uda antenna and obtained better results than other optimization techniques. In [14], Li has optimized Yagi-Uda antenna using Differential Evolution (DE) and illustrated the capabilities of the proposed method with several Yagi-Uda antenna designs. In [15], Singh *et al.* have analyzed another useful, stochastic global search and optimization technique known as Simulated Annealing (SA) for the optimization of Yagi-Uda antenna. In 2008, Dan Simon introduced a new optimization technique based on science of biogeography, in which information sharing among various habitats, i.e., potential solutions, is obtained via migration operator and exploration of new features is done with mutation operator [23]. Singh *et al.* have presented BBO as a better optimization technique for Yagi-Uda antenna designs, as compared to other optimization techniques in [16]. In [24], Li has proposed the Bi-Swarm optimization technique to optimize the Yagi-Uda antenna and produced better result than GA, Particle Swarm Optimization (PSO) and Computer Intelligence (CI) techniques. In 2011, Amaral *et al.* has applied Ellipsoid algorithm to optimization of Yagi-Uda antenna for gain maximization [25]. Li *et al.* in [26] have used Invasive Weed Optimization (IWO) technique to optimize a six element Yagi-Uda antenna for maximum directivity.

In this paper, 15 element Yagi-Uda antenna has been optimized for maximum gain using BBO and results are compared with other optimization techniques. A method of moments based freeware programme, Numerical Electromagnetics Code 2 (NEC2), is used to evaluate the antenna designs for gain.

After this brief introduction, the paper is structured as follows: In Section II, Yagi-Uda antenna is briefly discussed. Section III is dedicated to biogeography terminology and BBO technique. In Section IV, the design problem of 15 element Yagi-Uda antenna for gain maximization is presented and obtained results are compared with other optimization techniques. Finally, paper is concluded in Section V.

II. Antenna Design Parameters

Yagi-Uda antenna is basically made of three types of elements: (a) Reflector (b) Feeder and (c) Directors. *Reflector* is longest of all elements and blocks radiations in one direction. *Feeder* or driven element is fed with the signal to be trans-mitted, directly from transmission line. *Directors* are usually more than one in number and are responsible for unidirectional radiations. Normally, there is no limit on number of directors, however, as the number of directors are increased beyond a certain limit there is a reduction in the induced current in the most extreme elements. Figure 1 presents a basic Yagi-Uda antenna design where all elements are placed along y-axis and parallel to x-axis. Middle segment of the reflector is placed at origin and signal to be transmitted is fed to the middle segment of the feeder element. An incoming field induces resonant currents on all the antenna elements which causes parasitic (reflector and directors) elements to re-radiate signals. These re-radiated fields are then picked up by the feeder element, that makes total current induced in the feeder equivalent to combination of the direct field input and the re-radiated contributions from the director and reflector elements.

Element lengths and spacings between them are the variables/parameters which need to be determined for optimum design of Yagi-Uda antenna. An antenna with N elements requires 2N-1 parameters, i.e., N wire lengths and N-1 spacings, to be determined. These 2N-1 parameters, collectively, are represented as an integer vector referred as a *habitat* in BBO given as (1).

$$H= [L_1, L_2, . . . , L_N, S_1, S_2, . . . , S_{N-1}] \tag{1}$$

$$\alpha \ + \beta \ = \chi. \tag{1}$$

where L_N are the lengths and $S_{N-1}$are the spacings between antenna elements.

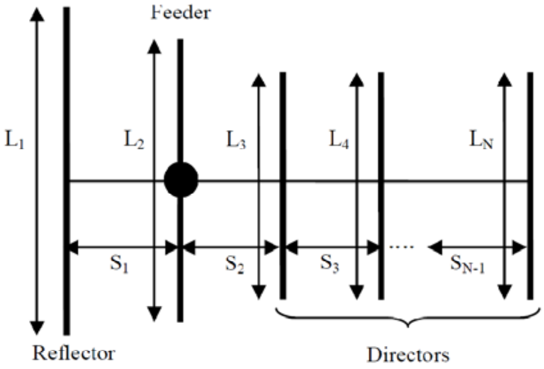


Fig. 1 Basic Yagi-Uda Antenna Design

III. Biogeography Based Optimization

Biogeography Based Optimization is a population based global optimization technique based on the science of bio-geography, i.e., study of the distribution of animals and plants among different habitats over time and space. BBO results presented by researches, to optimize Yagi-Uda antenna, are better than other optimization techniques like PSO, GAs, SA, DE etc. [10], [21], [13], [27].

Initially, biogeography was studied by Alfred Wallace [28] and Charles Darwin [29] mainly as descriptive study. However, in 1967, the work carried out by MacAurthur and Wilson [30] changed this perception by introducing a mathematical model for biogeography which made it possible to predict the number of species in a habitat. Mathematical models of biogeography describe the migration, speciation and extinction of species in various habitats.

A *habitat* or *island* is an ecological area inhabited by a particular animal species which is geographically isolated from other habitats. Each habitat is characterized by its Habitat Suitability Index (HSI). Habitats which are well suited as living places for biological species are referred to have high HSI value. HSI is analogues to fitness in other Evolutionary Algorithms whose value is a function of many features of the habitat such as rainfall, diversity of vegetation, diversity of topographic features, land area, and temperature etc. The features/variables that characterize habitability are known as Suitability Index Variables (SIVs). In other words, HSI is dependent variable whereas SIVs are independent variables.

The habitats with high HSI have large probability of emigration (hence high emigration rate, μ) simply due to large number of species they host and small probability of immigration (low immigration rate, λ) as they are already saturated with species. Immigration can be defined as the arrival of new species into a habitat, while emigration is the process of leaving one's native habitat. Similarly, habitats with low HSI tend to have low emigration rate, μ, due to sparse population, however, they will have high immigration rate, λ. Suitability of habitats having low HSI value is likely to increase with more number of species arriving from habitats having high HSI as suitability of a habitat depends upon its biological diversity. For sake of simplicity, it is safe to assume a linear relationship between HSI (or population) and immigration and emigration rates. Also maximum emigration and immigration rates are assumed equal, i.e., $E = I$, as shown graphically in Fig. 2.

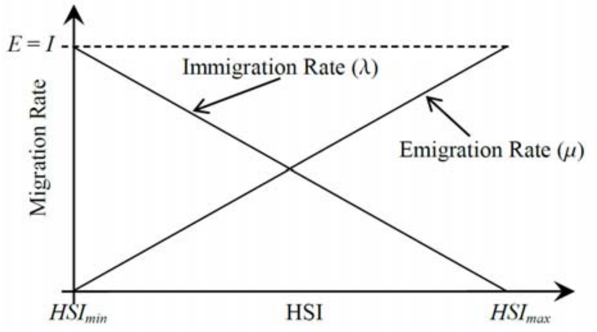


Fig. 2 Migration Curves

For k-th habitat, values of emigration rate, μ_k, and immigration rate, λ_k, are given by (2) and (3).

$$\mu_k = E.\frac{HSI_k}{HSI_{max} - HSI_{min}} \tag{2}$$

$$\lambda_k = I.\left(1 - \frac{HSI_k}{HSI_{max} - HSI_{min}}\right) \tag{3}$$

$$\alpha + \beta = \chi. \tag{1}$$

Good solutions (habitats with high HSI) are more resistant to change than poor solutions (habitats with low HSI) whereas poor solutions are more dynamic in nature and accept a lot of new features from good solutions. This addition of new features to low HSI solutions from high HSI solutions may raise the quality of those solutions.

In a global optimization problem with number of possible solutions, each habitat or a solution in a population of size *NP* is represented by *M*-dimensional integer vector as $H = [SIV_1, SIV_2, \ldots, SIV_M]$ where M is the number of SIVs (features) to be evolved for optimal HSI. HSI is the fitness criteria that is determined by evaluating the cost/objective function, i.e., HSI $= f$(H). BBO consists of mainly two mechanisms: (A) Migration and (B) Mutation, these are discussed in the following subsections.

A. Migration

Migration is a probabilistic operator that improves HSI of poor habitats by sharing information from good habitats. During migration, i-th habitat, H_i (where i = 1, 2, . . . , *NP*) use its immigration rate, λ_i given by (3), to probabilistically decide whether to immigrate or not. In case the habitat is selected for immigration, then the emigrating habitat, H_j, is found probabilistically based on emigration rate, μ_j given by (2). The process of migration is then carried out by copying values of SIVs from $H_j$to H_i randomly, i.e., $H_i(SIV) \leftarrow H_j(SIV)$. The migration process is depicted in Algorithm 1.

Algorithm 1 Standard Pseudo Code for Migration

```
fori = 1 to NP do
    Select $H_i$ with probability based on $\lambda i$
    if$H_i$ is selected then
for$_j$ = 1 to NP do
        Select $Hj$ with probability based on $\mu_j$
        if$H_j$ is selected
        Randomly select a SIV(s) from $H_j$
        Copy them SIV(s) in $H_i$
        end if
end for
end if
end for
```

B. Mutation

Mutation is another probabilistic operator that alters the values of randomly selected SIVs of some habitats that are intended for exploration of search space for better solutions by increasing the biological diversity in the population. Here, higher mutation rates are investigated on habitats those are, probabilistically, participating less in migration process. Elitism approach is usually used along with mutation to preserve features of the best habitat. The mutation rate, $mRate$, for k-th habitat is calculated as (4)

$$mRate_k = C \times \min(\mu_k, \lambda_k) \tag{4}$$

where $\mu_k$and λ_k are emigration and immigration rates, respectively, given by (2) and (3) corresponding to HSI_k. Here C is a scaling constant and its value is equal to 1. The pseudo code of mutation operator is given in Algorithm 2.

Algorithm 2 Standard Pseudo Code for Mutation

```
$mRate$ = $C$ × min($\mu_k$,$\lambda_k$), where $C$ = 1
**for**$i$ = 1 to $NP$ do
**for**$j$ = 1 to length($H$) do
        Select $H_j$(SIV) with $mRat$
        **if**$H_j$(SIV) is selected then
        Replace $H_j$(SIV) with randomly generated SIV
        **end if**
**end for**
**end for**
```

IV. Simulation Results and Discussions

Fifteen-wire Yagi-Uda antenna is optimized for maximum gain using BBO. To present a fair analysis, design is optimized with 30 habitats using 100 iterations. The C++ programming environment is used for development of optimization algorithm, whereas, a method of moments based software named as Numerical Electromagnetics Code 2 (NEC2) [31] used for determination of required antenna characteristic, i.e., gain. Each potential solution in BBO is encoded as an integer vector with 29 SIVs as given by (1). The radiation characteristics of Yagi-Uda antenna can change significantly by varying the element lengths and spacings up-to four decimal places, so this optimization algorithm finds the optimum element lengths and spacings between them. The search spaces for the search of optimum values of wire lengths and wire spacings are 0.30λ–0.50λ and 0.10λ–0.50λ, respectively. Cross sectional and segment sizes of all elements are kept constant, i.e., 0.003397λ and 0.1λ respectively, where λ is the wavelength corresponding to frequency of operation, i.e., 300 MHz. The scaling constant C, the maximum migration rates E and I, are set equal to 1. The corresponding lengths and spacings obtained during optimization of Yagi-Uda antenna with BBO are tabulated in Table I along with other optimization techniques from published work. It can be seen from the Table I that maximum gain of 18.41 dBi obtained with BBO is more than obtained by Bi-swarm optimization technique [24], Ellipsoid algorithm [25] and GA [10]. To the best of literature available, gain obtained by BBO, i.e., 18.41 dBi is the highest gain that is obtained from a 15-element Yagi-Uda antenna yet.

Table 1 Results of Gain Optimized 15 Element Yadi-Uda Anteena Designs

	BBO		**Bi-Swarm [24]**		**Ellipsoid [25]**		**GA [10]**	
Element	**Length**	**Spacing**	**Length**	**Spacing**	**Length**	**Spacing**	**Length**	**Spacing**
1(λ)	0.4808	-	0.4855	-	0.4561	-	0.474	-
2(λ)	0.4581	0.2690	0.4557	0.2397	0.4510	0.3788	0.486	0.356
3(λ)	0.4441	0.1963	0.4399	0.2810	0.4490	0.3921	0.452	0.144
4(λ)	0.4274	0. 3716	0.4311	0.3688	0.4430	0.4036	0.436	0.340
5(λ)	0.4154	0.4199	0.4259	0.3881	0.4375	0.4088	0.414	0.447
6(λ)	0.4061	0.4782	0.4215	0.3837	0.4445	0.4249	0.420	0.362
7(λ)	0.4067	0.4460	0.4043	0.4850	0.4473	0.4307	0.414	0.370
8(λ)	0.4029	0.4410	0.4054	0.4712	0.4431	0.4272	0.398	0.395
9(λ)	0.3970	0.4728	0.4033	0.4845	0.4373	0.4301	0.414	0.414
10(λ)	0.4017	0.4597	0.4094	0.4144	0.4386	0.4874	0.376	0.425
11(λ)	0.4013	0.4458	0.4028	0.4614	0.4229	0.3911	0.338	0.296
12(λ)	0.3991	0.4811	0.4074	0.4580	0.4213	0.4391	0.398	0.334
13(λ)	0.3991	0.4472	0.3936	0.5157	0.4385	0.3977	0.410	0.348
14(λ)	0.4015	0.4530	0.3955	0.4537	0.4867	0.4057	0.408	0.392
15(λ)	0.4146	0.4579	0.4142	0.4317	0.4293	0.2263	0.398	0.450
Gain (dBi)	**18.41**		**18.31**		**17.48**		**17.07**	

V. Conclusions and Future Scope

In this paper, optimization of fifteen-element Yagi-Uda antenna for gain maximization using BBO is carried out. As per observations, the gain obtained with BBO is higher as compared to other optimization techniques. The results show that BBO is a robust optimization technique for optimizing Yagi-Uda antenna. In the future scope of this paper, migration and mutation variants can be explored for better convergence performance.

References

[1] S. Uda and Y. Mushiake, *Yagi-Uda Antenna*, Maruzen, Ed. Maruzen Company, Ltd, 1954. [Online]. Available: http://books.google.co.in/ books?id=Uj9yYgEACAAJ

[2] H. Yagi, "Beam Transmission of Ultra Short Waves," *Proceedings of the Institute of Radio Engineers*, vol. 16, no. 6, pp. 715–740, 1928.

[3] D. G. Reid, "The Gain of an Idealized Yagi Array," *Journal of the Institution of Electrical Engineers-Part IIIA*: *Radiolocation*,, vol. 93, no. 3, pp. 564–566, 1946.

[4] J. Bojsen, H. Schjaer-Jacobsen, E. Nilsson, and J. Bach Andersen, "Maximum Gain of Yagi–Uda Arrays," *Electronics Letters*, vol. 7, no. 18, pp. 531–532, 1971.

[5] D. K. Cheng, "Optimization Techniques for Antenna Arrays," *Proceed-ings of the IEEE*, vol. 59, no. 12, pp. 1664–1674, 1971.

[6] L. C. Shen, "Directivity and Bandwidth of Single-band and Double-band Yagi Arrays," *IEEE Transactions on Antennas and Propagation*,, vol. 20, no. 6, pp. 778–780, 1972.

[7] D. Cheng and C. Chen, "Optimum Element Spacings for Yagi-Uda Arrays," *IEEE Transactions on Antennas and Propagation*,, vol. 21, no. 5, pp. 615–623, 1973.

[8] C. Chen and D. Cheng, "Optimum Element Lengths for Yagi-Uda Arrays," *IEEE Transactions on Antennas and Propagation*,, vol. 23, no. 1, pp. 8–15, 1975.

[9] D. K. Cheng, "Gain Optimization for Yagi-Uda Arrays," *Antennas and Propagation Magazine*, *IEEE*, vol. 33, no. 3, pp. 42–46, 1991.

[10] E. A. Jones and W. T. Joines, "Design of Yagi-Uda Antennas using Genetic Algorithms," *IEEE Transactions on Antennas and Propagation*,, vol. 45, no. 9, pp. 1386–1392, 1997.

[11] H. J. Wang, K. F. Man, C. H. Chan, and K. M. Luk, "Optimization of Yagi array by Hierarchical Genetic Algorithms," *IEEE*, vol. 1, pp. 91–94, 2003.

[12] N. Venkatarayalu and T. Ray, "Optimum Design of Yagi-Uda Antennas Using Computational Intelligence," *IEEE Transactions on Antennas and Propagation*,, vol. 52, no. 7, pp. 1811–1818, 2004.

[13] S. Baskar, A. Alphones, P. N. Suganthan, and J. J. Liang, "Design of Yagi-Uda Antennas using Comprehensive Learning Particle Swarm Optimisation," *IEEE*, vol. 152, no. 5, pp. 340–346, 2005.

[14] J. Y. Li, "Optimizing Design of Antenna using Differential Evolution," *IEEE*, vol. 1, pp. 1–4, 2007.

[15] U. Singh, M. Rattan, N. Singh, and M. S. Patterh, "Design of a Yagi-Uda Antenna by Simulated Annealing for Gain, Impedance and FBR," *IEEE*, vol. 1, pp. 974–979, 2007.

[16] U. Singh, H. Kumar, and T. S. Kamal, "Design of Yagi-Uda Antenna Using Biogeography Based Optimization," *IEEE Transactions on An-tennas and Propagation*,, vol. 58, no. 10, pp. 3375–3379, 2010.

[17] R. M. Fishenden and E. R. Wiblin, "Design of Yagi Aerials," *Proceed-ings of the IEE-Part III: Radio and Communication Engineering*, vol. 96, no. 39, p. 5, 1949.

[18] H. Ehrenspeck and H. Poehler, "A New Method for Obtaining Max-imum Gain from Yagi Antennas," *IRE Transactions on Antennas and Propagation*,, vol. 7, no. 4, pp. 379–386, 1959.

[19] E. Altshuler and D. Linden, "Wire-antenna Designs using Genetic Algorithms," *Antennas and Propagation Magazine, IEEE*, vol. 39, no. 2, pp. 33–43, 1997.

[20] D. Correia, A. J. M. Soares, and M. A. B. Terada, "Optimization of gain, impedance and bandwidth in Yagi-Uda Antennas using Genetic Algorithm," *IEEE*, vol. 1, pp. 41–44, 1999.

[21] N. V. Venkatarayalu and T. Ray, "Single and Multi-Objective Design of Yagi-Uda Antennas using Computational Intelligence," *IEEE*, vol. 2, pp. 1237–1242, 2003.

[22] Y. Kuwahara, "Multiobjective Optimization Design of Yagi-Uda An-tenna," *IEEE Transactions on Antennas and Propagation*,, vol. 53, no. 6, pp. 1984–1992, 2005.

[23] D. Simon, "Biogeography-based Optimization," *IEEE Transactions on Evolutionary Computation*,, vol. 12, no. 6, pp. 702–713, 2008.

[24] J. Li, "A bi-swarm optimizing strategy and its application of antenna design," *Journal of Electromagnetic Waves and Applications*, vol. 23, no. 14-15, pp. 1877–1886, 2009.

[25] A. Amaral, U. Resende, and E. Goncalves, "Yagi-uda antenna optimiza-tion by elipsoid algorithm," pp. 503–506, 2011.

[26] Y. Li, F. Yang, J. OuYang, and H. Zhou, "Yagi-uda antenna optimization based on invasive weed optimization method," *Electromagnetics*, vol. 31, no. 8, pp. 571–577, 2011.

[27] M. Rattan, M. S. Patterh, and B. S. Sohi, "Optimization of Yagi-Uda Antenna using Simulated Annealing," *Journal of Electromagnetic Waves and Applications*, 22, vol. 2, no. 3, pp. 291–299, 2008.

[28] A.Wallace, "The Geographical Distribution of Animals," *Boston, MA: Adamant Media Corporation*, vol. Two, pp. 232–237, 2005.

[29] C. Darwin, "The Orign of Species," *New York: gramercy*, vol. Two, pp. 398–403, 1995.

[30] R. MacArthur and E. Wilson, *The Theory of Island Biogeography*. Princeton Univ Pr, 1967.

[31] G. J. Burke and A. J. Poggio, "Numerical Electromagnetics Code (NEC) method of moments," *NOSC Tech. DocLawrence Livermore National Laboratory, Livermore, Calif, USA*, vol. 116, pp. 1–131, 1981.

A Review on Performance Comparison of Artificial Intelligence Techniques Used for Intrusion Detection

Navaneet Kumar Sinha[1], Gulshan Kumar[2] and Krishan Kumar[3]
[1]*Department of Computer Science & Engineering,*
Punjab Institute of Technology, Kapurthala, Punjab, India
[2]*Department of Computer Applications, SBS State Technical Campus, Ferozepur, Punjab, India*
[3]*Department of Computer Science & Engineering, SBS State Technical Campus, Ferozepur, Punjab, India*
E-mail:[1]*navaneetsinha@gmail.com,* [2]*gulshanahuja@gmail.com,* [3]*k.salujasbs@gmail.com*

***Abstract*—In the current era of Internet, network security technology has become crucial in protecting the computing infrastructure on the network. The number of network attacks has risen, leading to the essentials of network intrusion detection systems (IDS) to secure the network. Optimizing the performance of IDS becomes an important open problem which receives more and more attention from the research community. In this work, implementation of Artificial intelligence based techniques in IDS is popular in the research community. The network traffic can be classified into normal and anomalous in order to detect intrusions. There are several classification techniques available to detect the attacks. Researchers compare these techniques and try to identify the best techniques for the different attack category. This paper presents a review on comparison of these techniques.**

Keywords: Intrusion Detection, Artificial Intelligence, KDDCup, Data Mining Techniques, Classification

I. INTRODUCTION

This is the era of the Internet and information system in which computing infrastructure and communication resources are shared over the open world of the Internet. However, this inter connectivity between computers also enables malicious users to misuse resources and mount an Internet attack. The continuously growing Internet attacks pose service challenges to develop a flexible, adaptive security oriented methods. Intrusion Detection System (IDS) is one of the most important components being used to detect Internet attacks [1]. Intrusion Detection System (IDS) is placed inside the protected network, looking for potential threats in network traffic and or audit data recorded by a host.

IDS are split into two categories: misuse detection systems and anomaly detection systems. Misuse detection is used to identify intrusions that match known attack scenarios. However, anomaly detection is an attempt to search for malicious behaviour that deviates from established normal patterns [2].

In order to detect the intrusion, various approaches have been developed and proposed [1]. The major techniques are Statistics based IDS, the behaviour of the system is represented by a random view point. On the other hand, knowledge based IDS techniques try to capture the claimed behaviour from available system data (protocol specification, network traffic instances, etc.). AI based IDS techniques involves establishment of an explicit or implicit model that allows the patterns to be categorized. In this paper our interest is in AI based IDS techniques.

Many authors have divided AI based techniques into different classes [1] [3]: Decision tree based techniques, Rule based techniques, Data Mining techniques, machine learning techniques and clustering techniques. These techniques are further classifies in different categories. Major Data mining techniques are fuzzy logic and Genetic algorithm based techniques. Major Machine learning techniques are Neural network (NN), Bayesian network, Markov model, Support Vector machine (SVM) and Clustering techniques etc.

In many papers the above techniques are tested on a dataset. They analysed the performance of the technique and also compared some techniques in term of different attacks [4]. In most of the papers KDDCup99 [4] [5] datasets are used to perform the test of AI technique. Because the KDD data set is widely used by researchers.

The KDD cup 1999 dataset set is based on the DARPA98 dataset which was built by the Defense Advanced Research Projects Agency (DARPA) in 1998 during the DARPA98 IDS evaluation program. KDD99 dataset is openly available on [5]. Two types of files KDD Training set and KDD Test set are available for training and testing purpose. The dataset has 41 attributes and one class attribute. Various researchers have used different feature reduction techniques to select most relevant and ir-redundant features of a dataset of intrusion detection system [6]. Because, presence of irrelevant and redundant features degrades the accuracy of results and increases the computational overhead.

The dataset is categories in five classes, four attack classes and a normal. Attacks fall into four main categories [5]:

1. DOS: denial-of-service, e.g., syn flood.
2. R2L: unauthorized access from a remote machine, e.g., guessing password.
3. U2R: unauthorized access to local superuser (root) privileges, e.g., various “buffer overflow” attacks.

4. Probing: surveillance and other probing, e.g., port scanning.

In this paper, we present the performance comparison of different AI techniques (ANN, Classification, clustering, SVM etc.) performed by many researchers. We also compare their work on the basis of different criteria such techniques used for comparison, dataset used, metrics evaluated, best performance technique, advantage, disadvantage etc.

The rest of this paper is organized as follows: In section II Literature Survey is discussed. In section III we present a table of their work on the basis of different criteria. Finally, section IV shows the summary.

II. LITERATURE SURVEY

Mukkamala and Sung (2003) investigated and compared the performance of IDS based on support vector machines (SVM), artificial neural network (ANN), multivariate adaptive regression splines (MARS) and linear genetic programs (LGPs) [7]. For experiment they used DARPA dataset on 5-class classification. They perform experiments on two randomly generated separate datasets of size 5092 and 6890 for training and testing. Through the variety of experiments they found that, with appropriately chosen population size, program size, crossover rate and mutation rate, LGPs outperform other techniques in term of detection accuracy at the expense of time. They also conclude comparative performance between others. MARS is superior to SVMs in respect to classifying U2R and R2L attacks. SVMs outperform ANNs in respect of scalability, training and running time, and prediction accuracy. Resilient back propagation achieved the best performance among the neural networks in terms of accuracy and training. But performance comparisons are based on very least performance metrics (detection accuracy, training time and testing time).

Nguyen and choi (2008) evaluated the performance of a set of classifiers on KDD dataset and based on the result they choose best algorithm for each attack category [8]. They also proposed two classifier algorithm selection models. They performed experiments on weka machine learning tool and KDD99 dataset. Ten widely used classifier algorithms BayesNet, NaiveBayes, J48 (C4.5 Decision Tree), NBTree, Decision Table, JRip (RIPPER), OneR, MLP (Multilayer Perceptron), SMO and LBk are evaluated on the basis of four attacks categories (DoS, Probe, U2R and R2L). To compare these classifiers, they used TP (True Positive) and FP (False Positive) of each algorithm. They also measured AA (average accuracy) and TT (training time) performance metric.

The advantage of their work is the comparative analyses are based on attack categories. Because no single algorithm could detect all attack categories with high detection rate and low false alarm, the result shows that for a given attack category, certain algorithms demonstrate superior performance compared to others. The best algorithms for each attack categories are identified as: JRip for DoS and Probe, Decision table for U2R and OneR for R2L. On the basis of this result a parallel model for classifier selection (JRip, Decision Table and OneR) is proposed in this paper. They also proposed a model for real time application classifier selection (J48, BayesNet and OneR).

Sadoddin and Ghorbani (2007) conducted blind experiments of unsupervised techniques on KDD99 dataset to analyze the performance of unsupervised techniques considering their main design choice [9]. In this paper algorithms of the three categories are studied Clustering techniques, Unsupervised SVM and K-Nearest-Neighbor. Clustering techniques include K-means, C-means, EM, Self-organizing Map (SOM), Y-means and Improved Competitive Learning Network (ICLN). The evaluation of the algorithm in this paper is done with various distributions of training and testing datasets. To carry out experiment different tools for different algorithm are used, Fuzzy Clustering and Data Analysis Toolbox for C-means, SOM Toolbox for SOM, LIBSVM library for One-Class SVM and Weka tool for EM. For Clustering techniques two sets of experiment are performed. In the first set performance of each clustering technique evaluated with two labeling heuristics, count-based and distance-based. At second set of experiment, the performance of each clustering technique is evaluated in direct versus indirect mode. In the result they concluded that direct-based is on the average dominant over count-based heuristic in almost all of the clustering techniques. The clustering techniques (Except Y-means) in indirect mode, perform better when trained with Train_8020 (percentage of normal and attack records is 80% and 20%, respectively), while USVM and Y-means perform better when trained with Train_9604. In direct mode, the performance of KNN-based outlier detection schemes decreases as the population of attack data increases in the target dataset. They also highlighted two observations. First, all techniques perform poorly in detecting R2L attack. Secondly, USVM and Y-means are clearly superior over other techniques in detecting U2R attacks. Fuzzy clustering is not suitable for distinguishing normal and abnormal data in intrusion detection because C-means delivers the worst results in almost all experiments. In this paper only unsupervised techniques are discussed and on the basis of very few performance metrics detection rate, false alarm rate and ROC curve.

Kumar and Kumar (2011) performed a set of experiment of supervised classifiers on benchmarked KDD cup 1999 dataset [10]. They analyzed common supervised classifiers used in literature for intrusion detection. Performance of various AI techniques is compared from different categories viz: Rule based, Tree based, Functions, Lazy, Bayes and Meta. Kumar [4] identified. Some standard performance metrics are F-measure (FM), classification rate (CR), false positive rate (FPR), cost per example (CPE), precision (PR), root mean square error (RMSE), area under ROC curve (ROC), and detection rate (DR). The advantage of this research is that they first identified best classifiers for each attack class in the respective classifier category. Secondly, they compared best classifiers in the respective category to identify overall best classifier for different class attack. Because classifiers are designed by keeping in mind to optimize different criteria so it is very significant to compare classifiers in each classifier category. In this paper, it is concluded that bagged tree-J48 classifier is the best and the stable classifier with the overall correct classification of malicious traffic with minimum CPE, FPR and maximum ROC. It is also found that rule based JRip and Bagged tree-J48 for probe, Bagged tree-J48 for DoS, JRip for U2R and Naïve Bayes, bagged tree-J48 and neural network based MLP for R2L attack class can be better performed classifiers. They also reported that a single classifier cannot detect all the attack classes efficiently and suggested that a set of classifiers might be used to detect different attack classes. It is also observed from these experiments that all supervised classifiers are poor perform in detecting U2R and R2L attack classes.

Sabhnani and Serpen (2003) evaluated the performance of a comprehensive set of machine learning algorithms on four attack categories in the KDD 1999 cup dataset [11]. They selected nine algorithms from the variety of fields: neural networks, probabilistic models, statistical models, fuzzy-neuro system and decision tree. The algorithms identified are: Multilayer perceptron (MLP), Gaussia classifier (GAU), K-means clustering (K-M), Nearest cluster algorithm (NEA), Incremental radial basis function (IRBF), Leader algorithm (LEA), Hypersphere algorithm (HYP), Fuzzy ARTMAP (ART) and decision tree (C4.5). The classifiers are compared with the performance metric probability of detection (PD) and false alarm rate (FAR). The all classifiers tested on KDD data sets offered an acceptable level of misuse detection performance for only two attack classes Probe and DoS (poor for U2R and R2L). The results of the experiment show that for a given attack category, certain algorithms demonstrate superior detection performance compared to others. Finally, they concluded that MLP performs the best for probing, K-M for DoS as well as U2R, and GAU for R2L attack categories. On the basis of this conclusion sabhnani and Serpen [11] proposed a multi-classifier model which is able to perform best for all four attack classes Probe, DoS, U2R and R2L.

The multi-classifier model consists of different algorithms best for each attack category, as sub-classifiers: MLP for detection of probe attack, K-means for DoS as well as U2R attacks, and GAU for R2L attack.The Performance of this model is compared with KDD cup Winner, KDD Cup RunnerUp and Aggarwal and Joshi algorithms. The Multi-classifier model showed significant improvement in detection rate. The problem with this comparison is that the performance measured with very few performance metrics (PD, FAR and cost per example. And the other is the selection of classifiers to be compared was not follow standard.

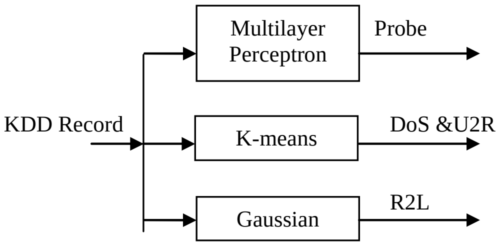


Fig. 1 Multi-Classifier Model

Wang *et al.* (2010) proposed a new neural network based detection approach called FC-ANN (fuzzy clustering based artificial neural network) [2]. The new FC-ANN approach is compared with selected well known classification approaches such as Decision tree, Naïve Bayes and BPNN. Performance metrics selected for this comparison are average accuracy, training time, precision, recall and F-value. The resulting analysis is critically done on several aspects. In terms of detection precision and detection stability FC-ANN outperforms BPNN and the other methods such as decision tree and NaïveBayes. In terms of average accuracy decision tree performs best and, also for probe and DoS attack class. Especially in case of low frequent attack classes U2R and R2L the new proposed approach FC-ANN gives significant improvement in detection precision and detection stability. The comparison is done for each attack class, but very less performance metrics are used.

Panda and Patra (2008) presented the comparison of three well known techniques such as ID3, J48 and Naïve Bayes [12]. The performance of classifiers is evaluated based on 10-fold cross validation test using KDD99 data set. The comparison is done with respect to performance metrics average accuracy, error rate, precision-recall, F-value, FPR, Area under the ROC curve, Kappa statistics and time taken to build the model. It is observed from all analysis that Naïve Bayes perform better than other two decision tree algorithms. However, decision trees are robust in detecting new intrusions, in comparison to the Naïve Bayes.

Kalyani and Lakshami (2012) presented the comparison of classification techniques such as Naive Bayes, J48, OneR, PART and RBF Network using NSL-KDD dataset [13]. The advantages of NSL-KDD dataset over KDDCUP’99 are also discussed. Several performance metrics are discussed such as TPR, FPR, RMSE, accuracy and time. J48 has higher accuracy, but they found PART as best algorithm because it takes lesser time, has lowest average error and accuracy is followed by J48.

Chauhan *et al.* (2013) presented the comparison of top ten classification algorithms: BayesNet, Logistic, SGD, IBK, JRip, PART, J48, Random Forest,Random Tree and REPT Tree [14]. To evaluate the algorithms 10-fold cross validation test is used. In experiment 20% of NSL-KDD data set is used and classifiers are tested on WEKA, a well known machine learning tool. The performance of all the classifiers is compared based upon accuracy, specificity and time. This study shows that decision tree classifiers are best at classifying the intrusions. Out of which Random Forest has outperformed with respect to the accuracy, specificity and sensitivity, whereas IBK consumes less time compared with others.

Gharibian and Ghorbani (2007) presented a comparison of supervised probabilistic and predictive machine learning techniques for intrusion detection [15]. Two probabilistic techniques NaiveBayes and Gaussian and two predictive techniques, Decision Tree and Random Forests are employed. In implementation, training data sets with different attack population and percentage are used to evaluate classifiers. Three different population categories used are 8020 (80% normal and 20% attack), 8416 and 8812. In the maximum detection rate analysis, Decision Trees and Random Forests show good results in detecting DoS, while Gaussian and NaiveBayes show better results in other attack categories. Other metrics analysis as sensitivity standard deviation and mean are also presented in this paper. Based on the results obtained in the paper [15], probabilistic techniques show more robustness than predictive techniques when trained using different training data sets. It has also been observed that probabilistic techniques show better detection rate in the data that has less samples such as R2L, U2R and Probe. While for DoS that has more samples, the predictive techniques outperform the probabilistic techniques.

Jalil *et al.* (2010) evaluated the performance of Decision tree (J48) classification algorithm and compared it with Support Vector machine (SVM) and Neural Network (NN) algorithms in term of accuracy, detection rate, false alarm rate and accuracy for four categories of attack under different percentage of normal data [16]. As summarized, from these four categories of attack (Probe, DoS, U2R, and R2L), Decision Tree (J48) has shown excellent results that outperform Neural Network and Support Vector Machines.

D’silva and Vora (2013) discussed three different clustering algorithms, namely K-Means Clustering, Y-Means Clustering and Fuzzy C-Means Clustering [17]. The comparison is made by taking into account various criteria like the performance, efficiency, detection rate, false positive rate, purity of cluster, etc. Among these Fuzzy C-Means clustering can be considered as an efficient algorithm for intrusion detection since it allows an item to belong to more than one cluster and also measures the quality of partitioning. Advantage and disadvantage of all three algorithms are discussed, but separate test for comparison purpose is not done.

Srinivasulu *et al.* (2009) also presented a comparison of widely used classification algorithms CART (Induction Decision Tree), Naïve Bayes and Artificial Neural network [18]. The test is performed on KDDCup99 data set in WEKA tool. All the three classifications, Dataset and WEKA tool are also discussed in brief. Performance is compared with TPR, FPR, Area under curve ROC, precision and recall, and all metrics are also discussed. The performance of the Induction tree (CART) method and ANN methods are better than the NB classifier. But the time taken is more for ANN than other classifiers.

Osareh and Shadgar (2008) compared the efficiency of machine learning methods in intrusion detection system, including artificial neural network and support vector machine [19]. They compare the accuracy, detection rate, false alarm rate for 4 attack types. In comparison, the research applies different normal data proportion for training and test, finally get one average value, and expect to obtain more objective results. In this paper, it is found that SVM is superior to NN in detection; in false alarm rate and in accuracy for Probe, Dos and U2R and R2L attacks, while NN could outperform the SVM only in accuracy.

Reddy *et al.* (2011) also presented a survey of various data mining techniques that have been proposed towards the enhancement of IDSs [20]. They also discussed the various AI techniques used in brief and also mentioned the drawbacks of IDS.

MeeraGandhi *et al.* (2010) also evaluated the performance of a set of classifier algorithms of rules (JRIP, Decision Tabel, PART, and OneR) and trees (J48, RandomForest, REPTree, NBTree) [21]. The algorithms are evaluated on KDD dataset. To compare the classifiers, TP (True positive) and FP (False Positive), Prediction Accuracy and learning time to build the model in seconds for each algorithm are considered. The results indicate that the C4.5 decision tree Classifier J48 outperforms in prediction than Rules. PART classifier, the Computational Performance differs significantly.

Neelima *et al.* (2014) presented a survey of the various data mining techniques that have been proposed towards the enhancement of IDSs [22]. Different data mining techniques used in intrusion detection are discussed in this paper.

Singh and Bansal (2013) presented the comparison of Multilayer Perception, Radial Base Function, Logistic Regression and Voted Perception [23]. They concluded that Multilayer Perceptron feed forward neural network has highest classification accuracy and lowest error rate as compared to other neural classifier algorithm network.

TABLE I SURVEY OF COMPARISON OF AI TECHNIQUES

S. No.	Paper	AI Techniques	Performance Metrics	Dataset	Advantage/Disadvantage	Best Techniques
1	Mukkamala and Sung [7]	MARS, SVM, LGP, ANN (RP, SCG, OSS)	DR, Training and Testing Time	DARPA	Anomaly Detection. Very least performance metrics	LGPs outperforms in term of accuracy at the expense of time.
2	Nguyen and Choi [8]	BayesNet, Naïve Bayes, J48, NBTree, Decesion Table, Jrip, OneR, MLP, SMO, LBK	TP, FP, Average Accuracy, Training Time	KDD99	Proposed a New Model for Real Time	JRip for DoS and Probe, Decision table for U2R, OneR for R2L in term of DR
3	Sadoddin and Ghorbani [9]	Clustering-K means, C-means, EM, SOM, Y-mean & ICLN, USVM, KNN	DR, FPR, ROC Curve	KDD99	Different tools for different algorithms used	
4	Kumar and Kumar [10]	RForest, RTree, NBTree, J48, Simple CART, Jrip, Decision Tree, NaïveBayes, BayesNet, SMO, MLP, RBFNetwork, LibSVM, IB1, LBK, K-star, Bagging, Boosting & Random SubspaceTree.	CR, CPE (Cost Per Example), RMSE, Precision (PR), ROC, AvG FM, Avg. DR, FPR	KDD99	Comparative Analysis of Techniques in each category as well as comparison b/w best classifiers of each category.	Bagged tree-J48 for overall correct classification, JRip and Bagged tree-J48 for probe, Bagged tree-J48 for DoS, JRip for U2R, Naïvebayes, bagged tree-J48 and MLP for R2L.
5	Sabhnani and Gerphen [11]	MLP, GAU, K-Mean, NEA, RBF, LEA, HYP, Fuzzy ARTMAP, C4.5	Detection Rate, FAR, CPE	KDD99	Multiple simulation tools are used. Less metrics selected. Proposed a MultiClassifier Model.	MLP for probing, K-M for DoS as well as U2R, and GAU for R2L.
6	Wang *et al.* [2]	Decision Tree, Naïve Bayes, BPNN, FC-ANN (proposed Method)	Precision, Recall, F-value, Avg. Accuracy, Training time,	KDD99	Evaluation for each type of Attacks and proposed an ANN based Approach.	Decision tree, FC-ANN in term of precision and recall, FC-ANN perform better for U2R and R2L
7	Panda and Patra [12]	Decision Tee(ID3 and J48) and Naïve Bayes	Avg. Accuracy, error rate, PR, ROC Area, Kappa statistics and time, F-Value, FPR	KDD99	Only three Classifiers are compared. 10-cross validation test performed	Naïve Bayes. Decision trees are robust in detecting new intrusions
8	Chauhan *et al.* [14]	BayesNet, Logistic, SGD, IBK, JRip, PART,J48, Random Forest, Random Tree and REPT Tree	Accuracy, sensitivity, specificity and time	NSL-KDD	The different training set is not used 10-fold cross validation is performed	Random Forest
9	Gharibian and Ghorbani [15]	NaiveBayes, Gaussian, Decision Tree and Random Forests	Detection rate, RMSE, standard deviation	KDD99	very few metrics are selected only four techniques are compared	Naïve Bayes and GAU are for DoS. Decision Tree and Random Forests for other attacks
10	Jalil *et al.* [16]	Detection Tree(J48), Support vector machine (SVM) and Neural Network (NN)	Avg. Accuracy, DR, FAR and accuracy for four attack classes.	KDD99	Different percentages of normal data are used. Performance in term attack classes. But only three techniques are compared.	Decision Tree (J48)
11	D’silva and Vora [17]	K-Means, Y-Means and Fuzzy C-Means Clustering	Efficiency, DR, FPR, purity of cluster	NA	Only clustering techniques are compared. No test result presented	Fuzzy C-Means
12	Srinivasulu *et al.* [18]	CART (Induction Decision Tree), Naïve Bayes and Artificial Neural network	TPR, FPR, F-measure, ROC Area, precision and recall	KDD99	Only three of classifiers are compared. Performance is not measured in terms of 4 attack categories	CART and Naïve Bayes
13	Osareh and Shadgar [19]	ANN and SVM	accuracy, DR, FAR	KDD99	Only ANN Technique and SVM is compared	SVM best in detection
14	MeeraGandhi *et al.* [21]	JRIP, Decision Tabel, PART, OneR, J48, RandomForest, REPTree, NBTree	TP, FP, Prediction Accuracy and Time to build the model	KDD99	rules and tree based approach are compared	C4.5 (J48)
15	Singh and Bansal [23]	RBF Network, Voted perceptron, Logistic Regression, Multilayer perceptron	CCI, ICI, KAPPA STATISTiCS, MAE, RMSE, RAE, RRSE, Time	NSL-KDD	Only ANN Techniques	Multilayer Perceptron

III. SUMMARY

The security is the primary concern in every field such as to prevent data from attacks and detect intruder. This paper has presented a survey of comparison of the various AI techniques that have been proposed towards the enhancement of IDSs. We presented literatures from the various papers and from the literature survey, it is analyzed that no single classification technique is sufficient to detect all four attack categories. Some researchers purposed to use the Multi classifier model to better perform for all attack classes. In most of the paper very few selected type of techniques are compared, there should follow a standard selection of techniques for comparing the performance. The NSL-KDD dataset has advantage over KDD99 but more researches are used KDD99 dataset.

REFERENCES

[1] Kumar *et al*.. "The use of artificial intelligence based techniques for intrusion detection: a review." *Artificial Intelligence Review*, vol. 34, no. 4, pp. 369-387, Dec. 2010.

[2] Wang *et al*. "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering." *Expert Systems with Applications,*vol. 37, no. 9, pp. 6225-6232, Sept. 2010.

[3] Chauhan *et al*. "Survey on data mining techniques in intrusion detection." *International Journal of Scientific & Engineering Research.vol.2 no. 7* ,July 2011.

[4] Tavallaee *et al*. "A detailed analysis of the KDD CUP 99 data set," *In Proceedings of the second IEEE Symposium on Computational Intelligence for Security and Defense Applications,* CISDA . pp. 1-6, 2009.

[5] "KDDCup 1999 Dataset". [Available online]: http://kdd.ics.uci.edu/databases/kddcup1999.html/

[6] Ahuja *et al*. "An empirical comparative analysis of feature reduction methods for intrusion detection." *International Journal of Information and Telecommunication Technology (ISSN: 0976-5972)* 1, no. 1, 2010.

[7] Mukkamala and Sung. "A comparative study of techniques for intrusion detection." In *Tools with Artificial Intelligence, Proceedings in 15th IEEE International Conference on*, pp. 570-577. IEEE, 2003.

[8] Nguyen and Choi. "Application of data mining to network intrusion detection: classifier selection model." In *Challenges for Next Generation Network Operations and Service Management*, pp. 399-408. Springer Berlin Heidelberg, 2008.

[9] Sadoddin and Ghorbani. "A comparative study of unsupervised machine learning and data mining techniques for intrusion detection." In *Machine Learning and Data Mining in Pattern Recognition*, pp. 404-418. Springer Berlin Heidelberg, 2007.

[10] Kumar and Kumar. "AI based supervised classifiers: an analysis for intrusion detection." In *Proceedings of the International Conference on Advances in Computing and Artificial Intelligence*, pp. 170-174. ACM, 2011.

[11] Sabhnani and Serpen. "Application of Machine Learning Algorithms to KDD Intrusion Detection Dataset within Misuse Detection Context." In *MLMTA*, pp. 209-215. 2003.

[12] Panda and Patra. "A comparative study of data mining algorithms for network intrusion detection." In *Emerging Trends in Engineering and Technology, ICETET'08. First International Conference on*, pp. 504-507. IEEE, July 2008.

[13] Kalyani and Lakshmi. "Performance assessment of different classification techniques for intrusion detection." *IOSR Journal of Computer Engineering, vol. 7* 2, no. 5, Nov. 2012

[14] Chauhan *et al*. "A Comparative Study of Classification Techniques for Intrusion Detection." In *Computational and Business Intelligence (ISCBI), 2013 International Symposium on*, pp. 40-43. IEEE, 2013.

[15] Gharibian and Ghorbani. "Comparative study of supervised machine learning techniques for intrusion detection." In *Communication Networks and Services Research, 2007. CNSR'07. Fifth Annual Conference on*, pp. 350-358. IEEE, 2007.

[16] Jalil *et al*. "Comparison of machine learning algorithms performance in detecting network intrusion." In *Networking and Information Technology (ICNIT), 2010 International Conference on*, pp. 221-226. IEEE, 2010.

[17] D'silva and Vora. "Comparative Study of Data Mining Techniques to Enhance Intrusion Detection." *International Journal of Engineering Research and Applications (IJERA) ISSN* pp. 2248-9622, 2013.

[18] Srinivasulu *et al*. "Classifying the network intrusion attacks using data mining classification methods and their performance comparison." *International Journal of Computer Science and Network Security* vol. 9, no. 6 pp. 11-18, 2009.

[19] Osareh and Shadgar. "Intrusion detection in computer networks based on machine learning algorithms." *International Journal of Computer Science and Network Security (IJCSNS)* vol. 8, no. 11, pp. 15-23, 2008.

[20] Reddy *et al*. "A study of intrusion detection in data mining." In *World Congress on Engineering (WCE),*vol. 3, pp. 6-8. 2011.

[21] MeeraGandhi *et al*. "Effective network intrusion detection using classifiers decision trees and decision rules." *Int. J. Advanced network and application, Vol. 2,* 2010.

[22] Neelima *et al*. "Leverage Data Mining Techniques in Intrusion detection.", *International Journal of Emerging Technology and Advanced Engineering, vol.4, pp.2 2014.*

[23] Singh and Bansal. "A Survey on Intrusion Detection System in Data Mining." *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET,). Vol. 2, no. 6, June 2013*

Performance Comparison of Wavelet Based and Conventional OFDM Systems – A Review

Harleen Kaur[1] and Gurpreet Singh[2]
[1,2]Department of Electronics & Comm. Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab India
E-mail: [1]harleen687@gmail.com, [2]er_gsandhu@yahoo.co.in

***Abstract*—Growth in technology has led to unprecedented demand for high speed architectures for complex signal processing applications. In 4G wireless communication systems, bandwidth is a precious commodity, and service providers continuously met with the challenge of accommodating more users with in a limited allocated bandwidth. To increase data rate of wireless medium with higher performance, OFDM (orthogonal frequency division multiplexing) is used. OFDM is multicarrier modulation (MCM) technique which provides an efficient means to handle high speed data streams on a multipath fading environment that causes ISI. Normally OFDM is implemented using FFT and IFFT's. To decrease the bandwidth waste brought by adding cyclic prefix, Wavelet based OFDM is employed. Due to use of wavelet transform the transmission power is reduced. In this paper we review the replacement of Fourier transform by Wavelet transform and its effects on the overall system.**

Keywords: BER, DFT, FFT, OFDM

I. Introduction

In the current and future mobile communications systems, data transmission at high bit rates is essential for many services such as high quality audio and video and mobile integrated service digital network. When the data is transmitted at high bit rates, over mobile radio channels, the channel impulse response can extend over many symbol periods, which lead to Inter-symbol Interference (ISI). In OFDM signal the bandwidth is divided into many narrow sub-channels which are transmitted in parallel [1] [2]. Each sub-channel is typically chosen narrow enough to eliminate the effect of delay spread. OFDM is multi-carrier modulation technique for transmission of signals over wireless channels. It converts a frequency-selective fading channel into a collection of parallel sub channels, which greatly simplifies the structure of the receiver. A combination of modulation and multiplexing constitutes to Orthogonal Frequency Division Multiplexing Independent signals that are sub-set of a main signal are multiplexed in OFDM and also the signal itself is first split into independent channels, modulated by data and then remultiplexed to create OFDM carrier. Orthogonality of subcarriers is the main concept in OFDM. The property of orthogonality allows simultaneous transmission of a lot of sub-carriers in a tight frequency space without interference from each other. This acts as an undue advantage in OFDM. Therefore, OFDM is becoming the chosen modulation technique for wireless communication. With the help of OFDM, sufficient robustness can be achieved to provide large data rates to radio channel impairments. In an OFDM scheme, a large number of orthogonal, overlapping narrow band sub-channels or sub-carriers transmitted in parallel by dividing the available transmission bandwidth. Compact spectral utilization with utmost efficiency is achieved with the help of minimally separated sub-carriers. Main attraction of OFDM lies with how the system handles the multipath interference at the receiver end. OFDM is multicarrier modulation (MCM) technique [3] which provides an efficient means to handle high speed data streams on a multipath fading environment that causes ISI. The spectral containment of the channels is better since it does not use cyclic prefix. One type of wavelet transform is Discrete Wavelet transforms have been considered as alternative platforms for replacing IFFT and FFT. [4] [5] [6]. It employs Low Pass Filter (LPF) and High Pass Filter (HPF) operating as Quadrature Mirror Filters satisfying perfect reconstruction and orthonormal properties. Wavelet transform [7–11] is a tool for studying signals in the joint time–frequency domains. Wavelets have compact support (localization) both in time and frequency domain, and possess better orthogonality. Orthogonal wavelets are capable of reducing the power of inter symbol interference (ISI) and inter carrier interference (ICI) which are caused by loss of orthogonality between the carriers as a result of multipath propagation over the wireless fading channels. In OFDM inter symbol interference (ISI) and inter channel interference (ICI) reduced by use of cyclic prefix (CP). In wavelet based OFDM, CP is not required. CP is 20% or more of symbol. Thus wavelet based OFDM gives 20% or more bandwidth efficiency [12–15]. Wavelet based OFDM is less affected by Doppler shift. In wavelet based OFDM a prototype wavelet filter provides both orthogonality and good time–frequency localization. Wavelet provides phase linearity and significant out-of-band rejection. Its energy compaction is also high. Wavelet based OFDM (WOFDM), which is also an MCM technique, possesses almost all advantages and disadvantages of conventional (Fourier based) OFDM. In this technique, the sub bandwidth division is obtained by using the inverse discrete wavelet based transforms, whereas conventional OFDM uses IFFT. Another main difference is that WOFDM symbols overlap in both time and frequency domains, whereas OFDM symbols overlap only in frequency domain. Therefore, adding CP to the WOFDM symbol frame does not have any effect on the bit error rate (BER) performance, as also shown in this work. One major advantage of WOFDM compared to OFDM is that WOFDM is more bandwidth efficient than OFDM.

In a wireless environment, the channel is much more unpredictable than a wire channel because of a combination of factors such as multi-path, frequency offset, timing offset, and noise. This results in random distortions in amplitude and phase of the received signal as it passes through the channel. These distortions change with time since the wireless channel response is time varying [16–18]. The channel estimation is a process of characterizing the effect of the transmission channel on the input signal. Channel estimation attempts to track the channel response. A dynamic estimation of channel is necessary before the demodulation of OFDM signals since the radio channel is frequency selective and time-variant for wideband mobile communication systems [19–20]. The channel estimate can then be used by an equalizer to correct the received constellation data so that they can be correctly demodulated to binary data.

This paper is organized as: in Section I, a brief introduction of the previous work demonstrated on COFDM and W-OFDM is discussed. The Section II presents the comparison of OFDM & WOFDM and result discussion, followed by the conclusion drawn in Section III on the basis of our observations.

II. OFDM & WOFDM

A. *Fourier-Based OFDM*

In OFDM, IFFT transform is used is used. In Fig. 1[21], the data {d$_k$} is processed by *M*-ary QAM modulator to map the data before IFFT, with *N* subcarriers. Its output is the sum of the information signals in the discrete time bearing as following:

$$x(k) = 1/\mathrm{N}\sum_{m=0}^{N-1}\left(X_{m^{e^{j2\pi km/N}}}\right) \tag{1}$$

where $\{X_k|0 \le k \le N - 1\}$ is a sequence in the discrete time domain, $\{X_m|0 \le m \le N-1\}$ are complex numbers in discrete frequency domain. The cyclic prefix (CP) is added before transmission to minimize the inter-symbol interference.

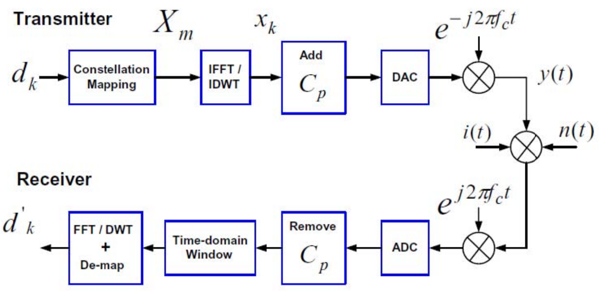


Fig. 1 OFDM Transceiver

At the receiver side, the processed is reversed to obtain and decoded the data. The CP is removed to obtain the data in discrete time domain. The data is then processed to the Time-Domain(TD) windowing for eliminating the narrowband interference before FFT. The output of FFT is the sum of the received signal in discrete frequency domain as follows:

$$X(m) = \sum_{k=0}^{N-1}\left(x_{k^{e^{-j2\pi km/N}}}\right) \tag{2}$$

B. *Wavelet-Based OFDM*

In the wavelet transform, inverse discrete wavelet transform (IDWT) and discrete wavelet transform (DWT) have replaced the IFFT and FFT in modulation and demodulation of FFT-OFDM system. Due to the overlapping nature of wavelet properties, the wavelet based does not need cyclic prefix to deal with delay spreads of the channel [22–24]. As a result, it has higher spectral containment than that of Fourier-based OFDM. The data {d$_k$} is processed as per FFT-OFDM. However, the difference is that the system does not require CP to be added to the OFDM symbol, and the system uses inverse discrete wavelet transform (IDWT) and discrete wavelet transform (DWT) to replace IFFT and FFT in transmitter and receiver, respectively. The output of the inverse discrete wavelet transform (IDWT) can be represented as:

$$s(k) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} S_m^n 2^{m/2}\psi(2_k^m - n)| \tag{3}$$

Where $\{S^n_m\}$ are the wavelet coefficients and $\psi(t)$ is the wavelet function with compressed factor *m* times and shifted n times for each subcarrier (number *k*, $0 \le k \le N-1$). The wavelet coefficients are the representation of signals in scale and position or time. The scale is related to the frequency. Low scale represents compressed wavelet which means that the signal is rapidly changing, or the signal is in high frequency. On the other hand, high scale represents stretched wavelet which means that the signal is slowly changing, or the signal is in low frequency. Thus, Xm can be represented to $\{S_m^n\}$ before it is processed to IDWT. At the receiver side, the process is inversed. The output of discrete wavelet transform (DWT) is

$$S_m^n = \sum_{k=0}^{N-1} s(k)2^{m/2}\psi(2_k^m - n) \tag{4}$$

S_m^n can be decoded to *Xm* before the recovery of data to QAM demodulator.

At first the results have been compared for the BER performance of Wavelet based OFDM system, and DFT based OFDM system over both Rayleigh fading and AWGN channel. The BER performance as a function of signal to noise ratio (Eb/No) is examined for Rayleigh fading and frequency selective fading with Doppler frequencies (fd = 200 Hz). Figs. 2 and 3 shows the BER performance of DFT based OFDM and wavelet based OFDM under Rayleigh fading channel and AWGN channel respectively.

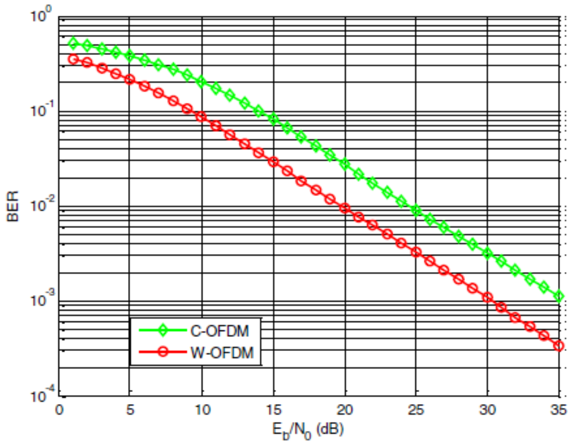


Fig. 2 [25] BER for C-OFDM and W-OFDM over Rayleigh Fading Channel

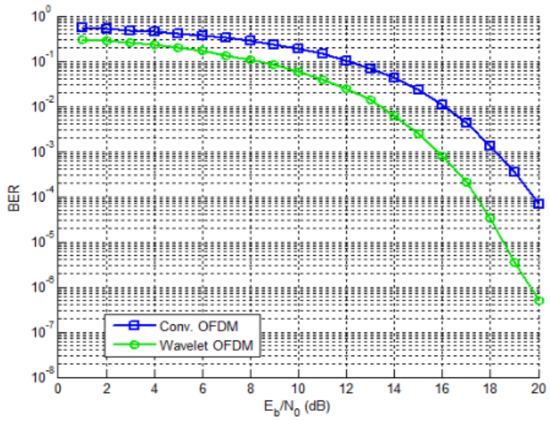


Fig. 3 [25] BER for C-OFDM and W-OFDM over AWGN Channel

It is clear from shown figures that Wavelet based OFDM have better BER performance as compared to DFT based OFDM.

Now the results have been evaluated for the MSE performance of Wavelet and DFT based OFDM system. For MSE performance, LMMSE estimator and LS estimator has been considered. Figures 3–4 show that LMMSE estimator has 15–20 dB better performance than LS estimator.

At last performance have been compared for the power spectral density (PSD) of wavelet based OFDM system with DFT based OFDM system. Fig. 5 shows that the wavelet based OFDM system is more spectral efficient then DFT based OFDM system. For comparative analysis between wavelet based OFDM system and conventional OFDM system, the Biorthogonal wavelet was considered. The Wavelet based OFDM system is out performing

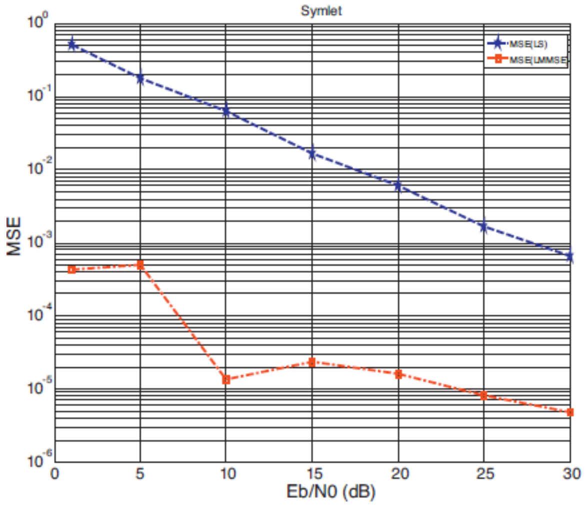


Fig. 3 [4] Performance of Wavelet Based OFDM System

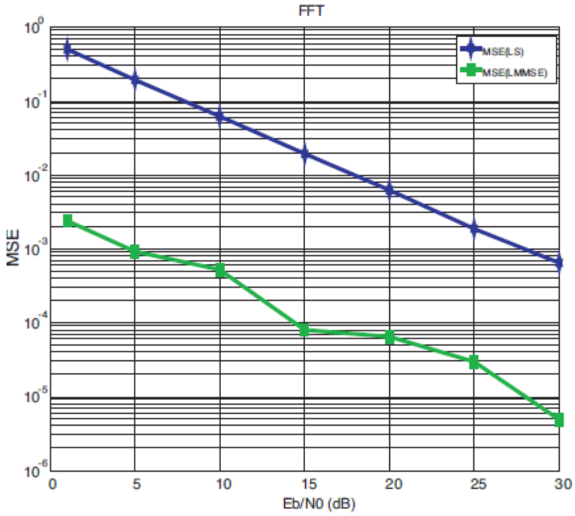


Fig. 4 [4] Performance of FFT Based OFDM System

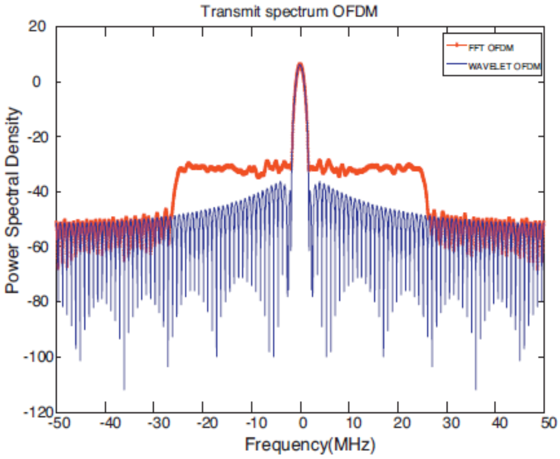


Fig. 5 [4] Comparison of Power Spectral Density

III. Conclusion

A comprehensive review of performance of conventional OFDM system and its comparison with wavelet based OFDM is presented in this work. Wavelet based OFDM system is a very flexible system which is also simple, and has a low complexity as only low order filters are needed instead of complex FFT processors and in addition, the filter type can be dynamically chosen depending on the condition of the channel or the data. In the comparison of BER performance of DFT and wavelet based OFDM system, wavelet based OFDM system have better performance in AWGN channel as well as Rayleigh fading channel. Wavelet based OFDM gives SNR improvement in AWGN channel As well as in Rayleigh fading channel. In the channel estimation of wavelet based OFDM and conventional OFDM the LMMSE estimator gives improvement than LS estimator. Power spectral density of wavelet based OFDM system is much better than conventional OFDM system. The main focus of this work is to put attention towards the realization of future high performance networks by introducing the Wavelet based OFDM in place of conventional OFDM systems.

References

[1] Mallat S., “A wavelet tour of signal processing”, 2nd edition, California: Academic Press, Elsevier, 1999.

[2] Lakshmanan M.K. and Nikookar H., “A review of wavelets for digital wireless communication”, Wireless Personal Communication, Vol. 37, pp: 387–420, 2006.

[3] R. Mirghani and M. Ghavami, “Comparison between Wavelet-based and Fourier-based Multicarrier UWB Systems”, IET Communications, Vol. 2, Issue 2, pp: 353-358, 2008.

[4] R. Dilmirghani and M. Ghavami, “Wavelet Vs Fourier Based UWB Systems”, 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp: 1-5, 2007.

[5] N. Ahmed, “Joint Detection Strategies for Orthogonal Frequency Division Multiplexing”, Dissertation for Master of Science, Rice University, Houston, Texas, pp: 1-51, 2000.

[6] S. D. Sandberg and M. A. Tzannes, “Overlapped Discrete Multi-tone Modulation for High Speed Copper Wire Communications”, IEEE Journal on selected areas in communications, Vol. 13, pp: 1571-1585, 1995.

[7] A. N. Akansu, and L. Xueming, “A Comparative Performance Evaluation of DMT (OFDM) and DWMT (DSBMT) Based DSL Communications Systems for Single and Multi-tone Interference”, proceedings of the IEEE International Conference on Acoustics, Speech and Signal processing, Vol. 6, pp: 3269-3272, 1998.

[8] Bingham JAC., “Multi-carrier modulation for data transmission: An idea whose time has come”, IEEE Communication Magazine, pp: 5-14, 1990.

[9] Resnikoff HL. and Wells RO., “Wavelet analysis”, Springer, 1998.

[10] Broughton SA., Bryan K., “Discrete Fourier analysis and wavelets”, New Jersey, John Wiley, 2009

[11] Pollet T., Van Bladel M. and Moeneclaey M., “BER sensitivity of OFDM systems to carrier Frequency offset and Wiener phase noise”, IEEE Trans Communication, Vol. 43, Issue 234, pp: 191-193, 1995.

[12] Chang S, Lee M, Park J. “A High speed VLSI architecture of discrete wavelet transform for MPEG-4”, IEEE Transactions on Consumer Electronics, Vol. 43, Issue 3, pp: 623-627, 1997.

[13] Mahesh Kumar Gupta and S. Tiwari, “Performance evaluation of conventional and wavelet based OFDM system”, Elsevier, 2012.

[14] Volkan Kumbasar, Oguz Kucur, “Performance comparison of wavelet based and conventional OFDM systems in multipath Rayleigh fading channels”, Digital Signal Processing, Vol. 22, pp: 841–846, 2012

[15] Abbas Hasan Kattoush, Waleed A. Mahmoudb, S. Nihad,” The performance of multiwavelets based OFDM system under different channel conditions”, Digital Signal Processing, Vol. 20, pp: 472–482, 2010

[16] M. Guatier, J. Lienard, and M. Arndt, “Efficient Wavelet Packet Modulation for Wireless Communication”, AICT’07 IEEE Computer Society, 2007.

[17] M. Guatier, and J. Lienard, “Performance of Complex Wavelet acket Based Multicarrier Transmission through Double Dispersive Channel”, NORSIG 06, IEEE Nordic Signal Processing Symposium (Iceland), June 2006.

[18] C. V. Bouwel, J. Potemans, S. schepers, B. Nauwelaers, and A. Van Caelle, “wavelet packet Based Multicarrier Modulation”, IEEE Communication and Vehicular Technology, SCVT 2000, pp. 131-138, 2000.

[19] C. Schurgers and M. B. Srivastava, “A systematic approach to peak – to – average power ratio in OFDM,” in SPIE’s 47th Annual meeting, San Diego, CA, 2001, pp. 454-464.

[20] Panchamkumar D SHUKLA, “Complex wavelet Transforms and Their Applications” Master Thesis 2003. Signal Processing Division. University of Strathclyde Department of Electronic and Electrical Engineering.

[21] Haixia Zhang, Dongfen Yuan, and Matthias Patzold, “Novel Study on PAPRs Reduction in Wavelet–Based Multicarrier Modulation Systems”, Elsevier, digital signal processing, 17(2007) 272-279, 5 sptember 2006.

[22] Zhou Lei, Li Jiandong, Liu Jing, and Zhang Guanghui, “A Novel wavelet Packet Division Multiplexing Based on Maximum Likelhood Algorithm and Optimal pilot Symbol Assisted Modulation for Reyleigh Fading Channels”, circuit system signal processing, vol. 24, No 3, 2005, PP. 287-302.

[23] G. Wornell, “Emerging Applications of Multirate Signal Processing and Wavelets in Digital Communications”, Proc. IEEE, Vol. 84, pp. 586–603, April 1996.

[24] N.G. Kingsbury, “The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters,” in Proc. 8th IEEE DSP Workshop, Utah.

[25] Vishal Sharma and Gurpreet Singh, “Wavelet-OFDM (W-OFDM) over Diverse Fading Channels” 2014 IEEE DOI 10.1109/ACCT.2014.43

Performance Analysis of MIMO Systems Using Equalizer and Combiner Techniques

Ekta Khurana[1] and Jaswinder Kaur[2]
[1,2]Department of Electronics & Comm. Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab, India
E-mail: [1]ektakhurana12@gmail.com, [2]jaswinder.ece@ gmail.com

***Abstract*—In this paper, performance of different MIMO systems is evaluated in terms of bit error rate (BER) using equalizer and combining technique. The combination of equalizer and combining technique (ZF-MRC and MMSE-MRC) is used to evaluate the BER performance for 2*2 MIMO systems using Binary phase shift keying (BPSK) and Quadrature phase shift keying (QPSK) modulation schemes for AWGN and Rayleigh channels. It is observed that ZF-MRC and MMSE-MRC performs superior with Binary phase shift keying (BPSK) for 2*2 MIMO systems in AWGN channel. Further it is found that MMSE-MRC gives better BER performance than ZF-MRC. In addition, 2*2 and 3*3 MIMO systems are compared using BPSK modulation scheme and AWGN channel. It is observed that 3*3 MIMO system performs better than 2*2 MIMO system.**

Keywords: MIMO, ZF, MMSE, MRC.

I. INTRODUCTION

Multiple-Input Multiple Output (MIMO) is a wireless technology that uses multiple antennas at both the transmitter and receiver to improve communication performance. It is an important part of modern wireless communication standards such as IEEE 802.11n (Wi-Fi), 4G, and WiMAX. It offers significant increase in data throughput and link range without additional bandwidth or transmit power [1] [8]. It is the basic building block of IEEE 802.11n standard. It provides better capacity and potential of improved reliability compared to single antenna channels. The combination of MIMO and OFDM (MIMO-OFDM) is a very effectual way to achieve high efficiency spectral wideband systems [7].

The equalizer technique is used to mitigate the effect of inter symbol interference (ISI) and combining technique is used to remove the fading. In this paper, different equalization and combiner techniques are used for the analysis of bit-error rate (BER) in different MIMO systems using different modulation schemes in Additive white gaussian noise and Rayleigh channel.

The Outline of this paper is as follows. Section II describes the MIMO system model. In Section III equalization te chniques are discussed. Section IV describes the diversity combining signal techniques. Section V includes the simulation and results. Conclusion is given in Section VI.

II. MIMO SYSTEM MODEL

A typical MIMO system with a transmit array of Tx_n antennas and a receive array of Rx_n antennas is shown in Fig. 1. These systems can utilize multipath components during transmission to solve the problem of multipath fading.

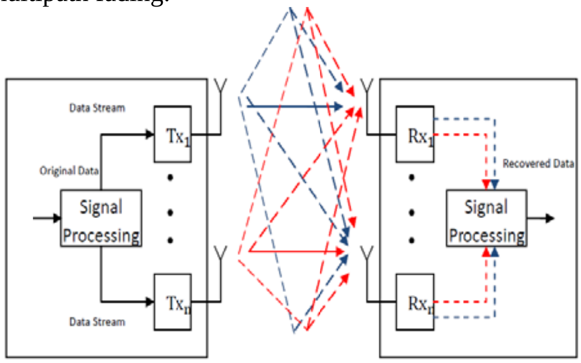


Fig. 1 MIMO System Model

In mathematical terms, MIMO is defined as [3]:

$$y = Hx + n \tag{1}$$

where y is the received signal, H is the matrix of transmitting and receiving antennas, x is the transmitted signal, n is the noise.

MIMO system provides better quality of service, low bit error rate and improved coverage area. The cost of these systems increases with increase in number of antennas.

III. EQUALIZATION TECHNIQUES

Equalization is a well known technique used by communication engineers to mitigate the effects of inter symbol interference (ISI) [2]. There are two types of equalization techniques-Linear and Nonlinear. Linear techniques are Zero forcing (ZF), Minimum mean square error (MMSE), etc. Non linear techniques are Decision feedback equalizer (DFE), Maximum likelihood sequence estimation (MLSE) etc.

A. Zero Forcing Equalizer

It is a form of linear equalization algorithm which was developed by Robert Lucky. To restore the signal, the inverse of the channel frequency response is applied to the received signal. It brings down the ISI to zero in noise free case. It gives poor performance of BER as it does not take into account the noise [3]. This algorithm only minimizes the ISI and peak distortion.

To find x in equation “1”, there is need to find a matrix W which satisfies WH = 1. The zero forcing equalizer meeting this constraint is given by [4]:

$$W_{ZF} = (H^H H)^{-1} H^H \tag{2}$$

where W_{ZF} is the Equalization matrix and H is the channel matrix.

B. *Minimum Mean Square Equalizer*

It is also a form of linear equalization algorithm that minimizes both ISI and noise. It is based on the mean square error (MSE) criterion [9]. This algorithm reduces the variance of the error signal and gives better performance than zero forcing.

MMSE minimizes mean square error given by E [3]:

$$E[W_y - x][W_y - y]^H \tag{3}$$

On solving equation “3”, we get

$$W_{MMSE} = (H^H H + N_o I)^{-1} H^H \tag{4}$$

where W is the Equalization matrix, H is the channel matrix, N_o is the noise variance and I is the interference power.

IV. Diversity Signal Combining Techniques

In wireless communications, signal fading is caused by multipath-effect. Diversity combining techniques are used to compensate this signal fading [5]. There are various combining techniques-Selection combining (SC), Feedback or Scanning combining, Maximal ratio combining (MRC), Equal gain combining (EGC).

A. *Maximal Ratio Combining*

Maximal ratio combining technique gives the best reduction of fading and is used by the receiver with multiple antennas. In this technique, all the signals are combined in co-phased and weighted manner so as to have the highest achievable signal to noise ratio at the receiver at all the times. It is also known as ratio-squared combining and predetection combining. The range of a wireless system is improved by this technique [6].

V. Simulation and Results

A. *2*2 MIMO System with ZF-MRC*

This section presents the evaluation of bit error rate (BER) for different modulation schemes (Binary phase shift keying (BPSK) & (Quadrature phase shift keying (QPSK)) with ZF-MRC in 2*2 MIMO system for Additive white Gaussian noise (AWGN) and Rayleigh channels.

Figure 2 shows the results for MIMO system with ZF-MRC using BPSK & QPSK for Additive white gaussian noise channel. It is found that acceptable BER of 10^{-3} is achieved in case of BPSK at signal to noise ratio = 13 dB as compared to QPSK with signal to noise ratio = 15 dB which shows that ZF-MRC performs better with Binary phase shift keying scheme for Additive white gaussian noise channel.

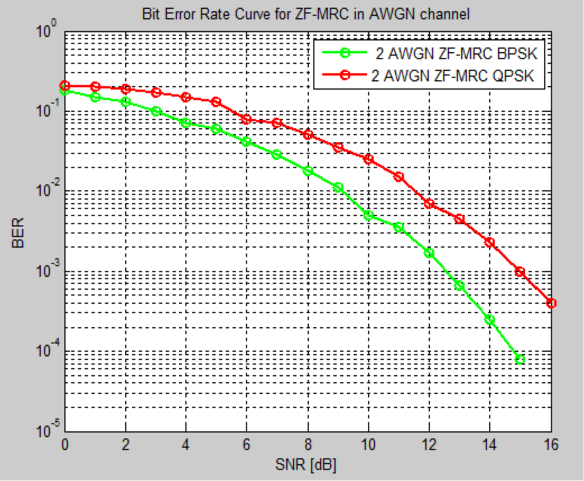


Fig. 2 2*2 MIMO System Using ZF-MRC in AWGN Channel

The results for system with ZF-MRC using Binary phase shift keying & Quadrature phase shift keying for Rayleigh channel is shown in Fig. 3. It is found that signal to noise ratio requirement to achieve the acceptable BER of 10^{-3} in case of BPSK is 15 dB which is smaller than QPSK signal to noise ratio requirement. This shows that ZF-MRC performs superior with Binary phase shift keying scheme for Rayleigh channel.

Figure 4 shows comparison of ZF-MRC for both the channels (Additive white gaussian noise (AWGN) & Rayleigh) with their best modulation scheme (Binary phase shift keying). It is found that ZF-MRC performs better for Additive white gaussian noise channel as the acceptable BER of 10^{-3} is achieved at signal to noise ratio = 13 dB in case of Additive white gaussian noise channel as compared to Rayleigh channel signal to noise ratio with 15 dB.

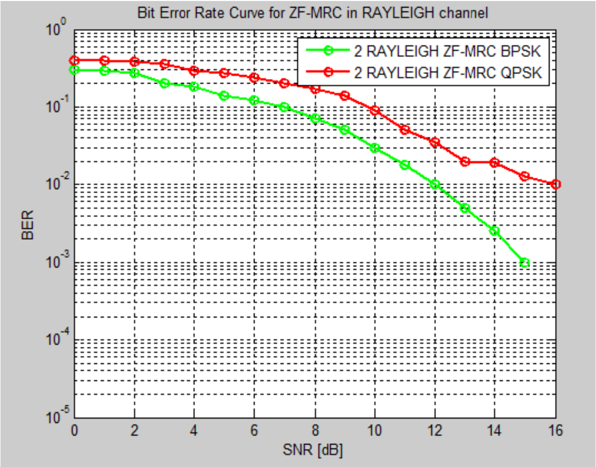


Fig. 3 2*2 MIMO System Using ZF-MRC in Rayleigh Channel

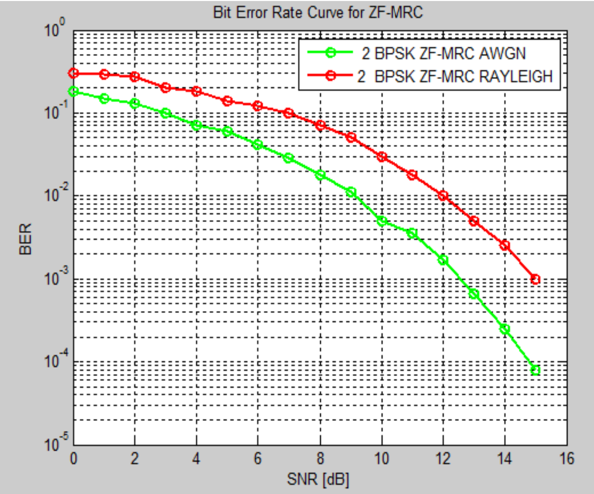


Fig. 4 2*2 MIMO System Using ZF-MRC in AWGN and Rayleigh Channel

*B. 2*2 MIMO System with MMSE-MRC*

This section presents the evaluation of bit error rate (BER) for different modulation schemes (Binary phase shift keying (BPSK) & (Quadrature phase shift keying (QPSK)) with MMSE-MRC in 2*2 MIMO system for Additive white Gaussian noise (AWGN) & Rayleigh channels.

The results for system with MMSE-MRC using Binary phase shift keying & Quadrature phase shift keying for AWGN channel is shown in Fig. 5. It is found that signal to noise ratio requirement to achieve the BER of 10^{-3} is 11 dB in case of BPSK which is smaller than QPSK signal to noise ratio requirement of 13 dB. This shows that MMSE-MRC performance is superior for AWGN channel with Binary phase shift keying scheme.

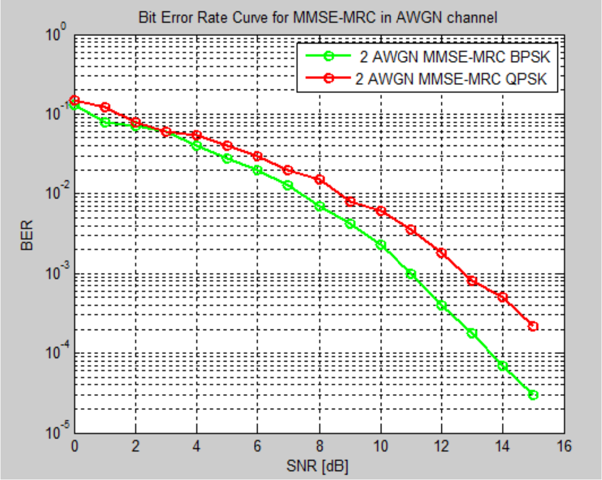


Fig. 5 2*2 MIMO System Using MMSE-MRC in AWGN Channel

Figure 6 shows the results for MIMO system with MMSE-MRC using BPSK & QPSK for Rayleigh channel. It is found that MMSE-MRC performs better with Binary phase shift keying scheme (BPSK) for Rayleigh channel as acceptable BER of 10^{-3} is achieved at signal to noise ratio = 14dB in case of BPSK which is less than QPSK signal to noise ratio requirement.

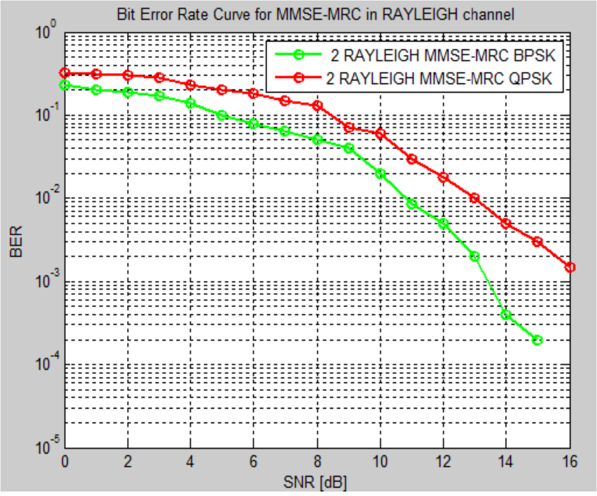


Fig. 6 2*2 MIMO System Using MMSE-MRC in Rayleigh Channel

The comparison of MMSE-MRC with best modulation scheme (Binary phase shift keying) for Additive white Gaussian noise and Rayleigh channel is shown in Fig. 7. It is observed that acceptable BER value of 10^{-3} is obtained at signal to noise ratio = 11 dB in case of Additive white Gaussian noise channel as compared to Rayleigh channel with signal to noise ratio = 14 dB. This shows that MMSE-MRC performs better for AWGN channel.

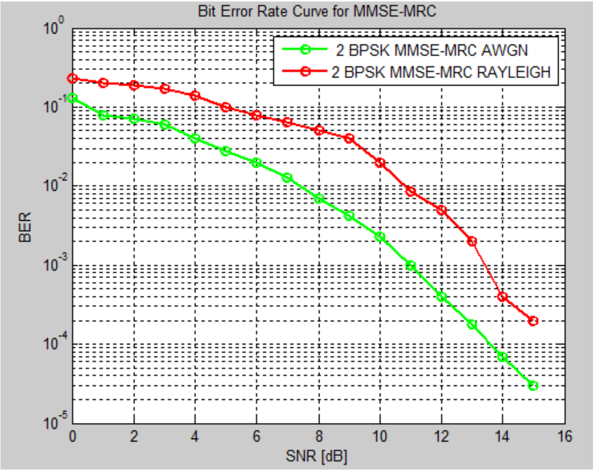


Fig. 7 2*2 MIMO System Using MMSE-MRC in AWGN and Rayleigh Channel

C. Comparison of ZF-MRC and MMSE-MRC

ZF-MRC and MMSE-MRC is compared with best modulation scheme (Binary phase shift keying) for Additive white Gaussian noise channel in Fig. 8. It is found that MMSE-MRC performs better than ZF-MRC as acceptable BER of 10^{-3} is achieved in case of MMSE-MRC at signal to noise ratio = 11dB as compared to ZF-MRC with signal to noise ratio = 13 dB.

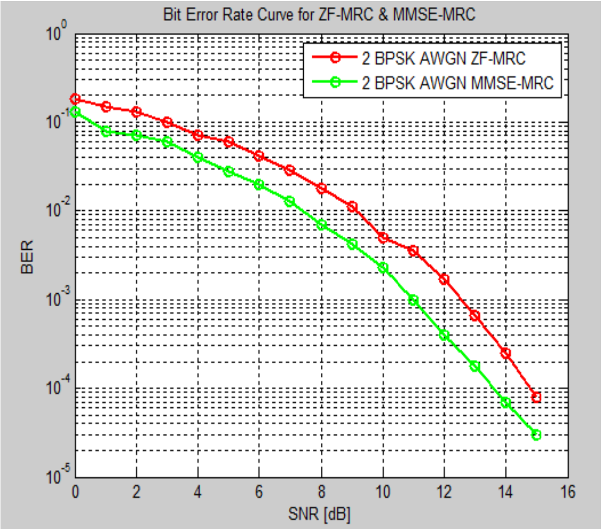


Fig. 8 ZF-MRC and MMSE-MRC Using AWGN Channel

Figure 9 shows the comparison of ZF-MRC AND MMSE-MRC with best modulation scheme (Binary phase shift keying) for Rayleigh channel. It is found that signal to noise ratio requirement to achieve the BER of 10^{-3} is 14 dB in case of MMSE-MRC which is smaller than ZF-MRC with signal to noise ratio = 15 dB. This shows MMSE-MRC again performs superior for Rayleigh channel.

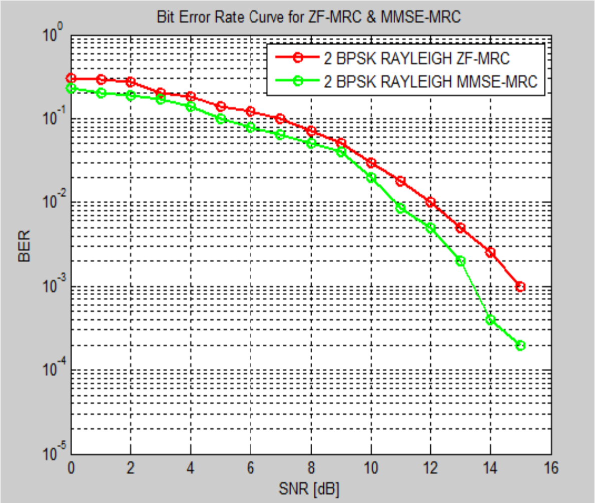


Fig. 9 ZF-MRC AND MMSE-MRC Using Rayleigh Channel

D. *Comparison of Different MIMO Systems Using ZF-MRC*

Different MIMO systems such as 2*2 and 3*3 are compared using ZF-MRC with Additive white Gaussian noise and best modulation scheme (Binary phase shift keying). It is found that signal to noise ratio requirement to achieve the acceptable BER of 10^{-3} in case of 3*3 MIMO system is 7 dB which is smaller than 2*2 MIMO system signal to noise ratio of 13 dB. This shows that 3*3 MIMO system performance is superior than 2*2 MIMO system.

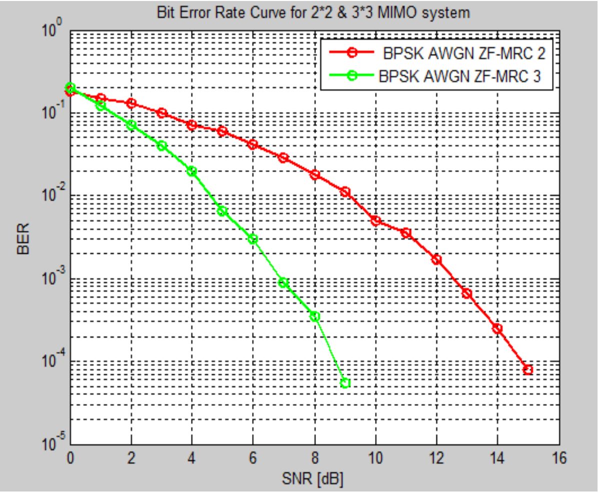


Fig. 10 2*2 and 3*3 MIMO Systems Using ZF-MRC

E. *Comparison of Different MIMO Systems Using MMSE-MRC*

In Figure 11 2*2 and 3*3 MIMO systems are compared using MMSE-MRC with Additive white Gaussian noise and best modulation scheme (Binary phase shift keying). In this case also, acceptable bit error rate of 10^{-3} dB is achieved by 3*3 MIMO system at signal to noise ratio = 7 dB as compared to 2*2 MIMO system with signal to noise ratio = 11 dB which shows that it is superior than 2*2 MIMO system.

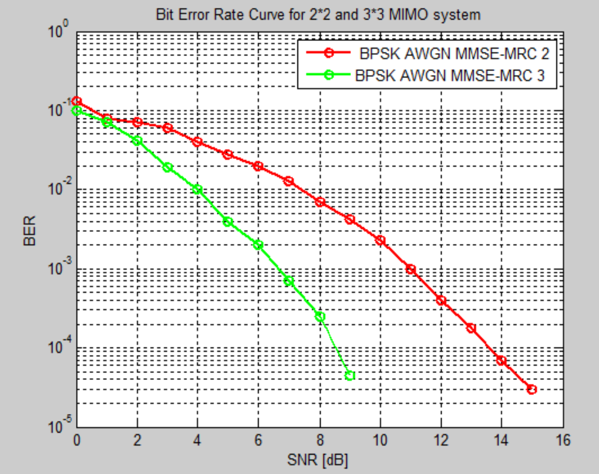


Fig. 11 2*2 and 3*3 MIMO Systems Using MMSE-MRC

VI. CONCLUSION

The BER performance is evaluated for different MIMO systems using equalizer and combining techniques with different modulation schemes for AWGN and Rayleigh channels. It is found that ZF-MRC and MMSE-MRC perform better in case of BPSK with AWGN and Rayleigh channel as it gives low BER as compared to QPSK. For both the channels, MMSE-MRC performs better than ZF-MRC in terms of BER. In addition, 2*2 & 3*3 are compared using best modulation scheme and best channel which shows that as the number of transmitting and receiving antenna increases, the bit error rate decreases. Future research will consist of evaluation of BER of different MIMO systems using nonlinear equalizer and combining techniques.

REFERENCES

[1] Tolga M. Duman, Ali Ghrayeb, “Coding for MIMO Communication Systems”, John Wiley & Sons, 2008.

[2] Bhasker Gupta, Davinder S. Saini, “BER performance improvement in OFDM systems using Equalization algorithms”, IEEE International Conference on Parallel Distributed and Grid Computing (PDGC), pp 49-54, 2010.

[3] Bhasker Gupta, Davinder S. Saini, “BER performance improvement in MIMO systems using various Equalization techniques”, IEEE International Conference on Parallel Distributed and Grid Computing (PDGC), pp 190-194, 2012.

[4] Rohit Gupta, Amit Grover, “BER Performance Ananlysis of MIMO Systems Using Equalization and Combing Techniques”, Innovative Systems Design and Engineering, Vol 3, pp 11-25, 2012.

[5] Theodore S. Rappaport, “WirelessCommunication: Principles and Practice”, 2nd ed., Prentice Hall, 2005.

[6] Gottapu Sasibhushana Rao, “Mobile Cellular Communication”, Pearson, pp 336-338.

[7] Nisha Achra, Garima Mathur, Prof R.P.Yadav, “Performance Analysis of MIMO OFDM System for Different Modulation Schemes under Flat Fading channels”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, pp 2098-2103, 2013.

[8] Arogyaswami J.Paulraj, Dhananjay A.Gore, Rohit U.Nabar, Helmut Bolcskei, “ An Overview of MIMO Communications–A Key to Gigabit Wireless”, Proceedings of the IEEE, Vol.92, pp 198–218, 2004.

[9] N.Sathish Kumar, Dr. K.R Shankar Kumar, “ Performance Analysis of M*N Equalizer based Minimum Mean Square Error (MMSE) Receiver for MIMO Wireless Channel”, International Journal of Computer Applications,Vol.16, pp 47-50, 2011.

A Semi Blind DWT-SVD Video Watermarking Under Attacks

Divjot Kaur Thind[1] and Sonika Jindal[2]
[1,2]Department of Computer Sc. & Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, India
E-mail: [1]prettythind01@gmail.com, [2]sonikamanoj@gmail.com

***Abstract*—Digital watermarking was introduced due to rapid advancement of networked multimedia systems. It was developed to enforce copyright technologies for protection of copyright ownership. This technology is first used for still images but recently they have been developed for other multimedia objects such as audio, video etc. In this paper a new digital video watermarking scheme is proposed which combines Discrete wavelet transform (DWT) and Singular Value Decomposition (SVD) in which watermarking is done in the high frequency sub band and then various attacks have been applied. Tests have been undergone to check the proposed scheme for robustness and imperceptibility.**

Keywords: Video Watermarking, Discrete Wavelet Transform (DWT), Singular Value Decomposition (SVD)

I. Introduction

The widespread expansion of the internet has led to availability of digital data such as audio, images and videos to the public which leads to the issue of data protection. Many techniques has been used but digital watermarking is new technique to issue these problems [1]. The watermarking is a covert security feature used for copyright protection, authentication. Digital watermarking means embedding the secret information in the form of watermark into the digital multimedia such as image, audio and video. The embedded information is extracted out to reveal the real owner/identity of the digital media. After embedding the watermark the original data should not alter. These watermarks should be robust against any kind of intended or unintended attacks. There are three parameters in digital watermarking: data payload, fidelity and robustness. Digital watermarking has been extensively used for still images but now they are used for other multimedia objects such as audio and videos [2]. Digital video watermarking is the process of embedding and extracting watermark from the videos. There are many algorithms of video watermarking some of which consider videos as group of continous still images. Some algorithms consider the temporal dimension. Watermarking techniques can be applied in two domains: Spatial domain and transform domain. Spatial domain technology embeds watermark directly into the pixels which changes the intensity values [3]. Previously watermarking techniques were based on spatial domain example least significant bits (LSBs). This method is easy and simple but affected by the attacks. Transform domain technology embeds watermark in the transform of the signal. These transform are discrete fourier transform, discrete wavelet transform (Chan and Lyu [4], [5]), discrete cosine transform (Cox et al [6], Hsu and Wu [7]). Section II introduces the two main concepts of the paper i.e., SVD and DWT. Section III introduces the proposed method of video watermarking. Section IV represents experimental results and section V gives the concluding remarks.

II. Preliminaries

A. Singular Value Decomposition (SVD)

The SVD is popular mathematical technique that provides tool for analysis of matrices. It is an elegant way for extracting algebraic features from an image [8]. It was first introduced by Beltrami and Jordan in 1870 for square matrices and then Eckart and Young in 1936 extended to rectangular matrices. SVD has provided its great application in image processing and watermarking. The SVD matrix of an image has good stability. When a small perturbation is added to an image, large variation of its SVs does not occur. Using this property of the SVD matrix of an image, the watermark can be embedded to this matrix without large variation in the obtained image. Let us consider an image A as matrix of size M*N. Using SVD matrix A can be decomposed as:

$$A = USV^T$$

$$U = [u_1, u_2, \dots\dots, u_n]$$

$$V = [v_1, v_2, \dots\dots\dots, v_n]$$

$$S = \begin{bmatrix} s_1 & . & . \\ . & s_2 & . \\ . & . & s_n \end{bmatrix}$$

where U and V are orthogonal matrices of size M*N and S is a diagonal matrix. The columns of V called right singular vectors and the columns of U are left singular vectors of the image A. In SVD based watermarking, SVD of the original image is taken and then singular values of the matrix are modified by introducing the singular values of watermark. The properties of svd that made it popular are as follows:

1. SVD do not affect the image quality.
2. It preserves non symmetric properties.
3. They are robust to various attacks such as rotation, scaling, compression, noise addition and cropping.
4. It extract algebraic properties of digital image.

B. *Discrete Wavelet Transform (DWT)*

Wavelet transform is time domain localized analysis method and it differentiates time in high frequency part of DWT transformed signals and frequency differentiated in low frequency parts of signals [9]. DWT is multiresolution mathematical tool for decomposing an image. An image is considered as two dimensional signal which when passed through high and low pass filters decompose into several sub bands having different resolutions. DWT decomposes an image into four components namely LL,HL,LH,HH where first letter corresponds frequency operation and second letter is the filter applied. LL represents approximate features of an image and it is half of the original image. LH (Vertical high frequency), HL (Horizontal high frequency) and HH (High frequency) represents detail of an image. It can further decompose by applying 2-level DWT on the sub-image. After applying a 2-level DWT, sub-image get decomposes into the approximation sub-band (LL2), the horizontal subband (LH2), the vertical sub-band (HL2), and the diagonal sub-band (HH2). Again decomposition of image results into LL3, LH3, HL3 and HH3 sub-band respectively. Several algorithms has been proposed on using dwt-svd, one such algorithm is proposed by Osama S. Faragallah [10]. It is an efficient and robust video watermarking technique based on SVD and DWT. In this technique, middle and high frequency bands are SVD transformed and watermark is hidden in that.

III. PROPOSED VIDEO WATERMARKING TECHNIQUE

A. *Video Watermark Embedding Process*

Figure 1 shows the algorithm of the watermark embedding process.

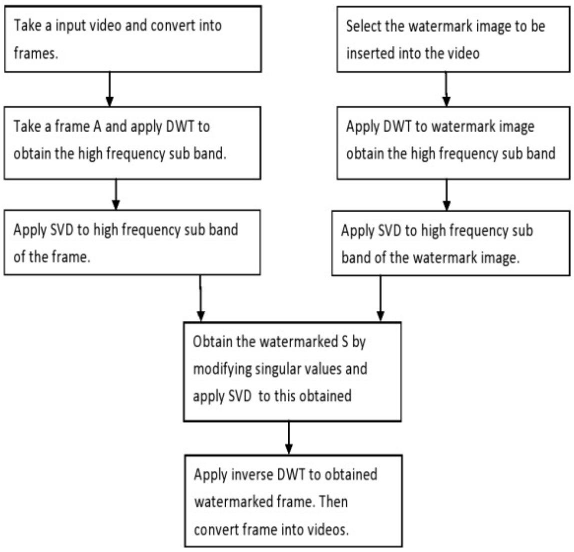

Take a input video and convert into frames.
Take a frame A and apply DWT to obtain the high frequency sub band.
Apply SVD to high frequency sub band of the frame.
Select the watermark image to be inserted into the video
Apply DWT to watermark image obtain the high frequency sub band
Apply SVD to high frequency sub band of the watermark image.
Obtain the watermarked S by modifying singular values and apply SVD to this obtained
Apply inverse DWT to obtained watermarked frame. Then convert frame into videos.

Fig. 1 Video Watermarking Embedding Process

B. *Video watermark Extraction Process*

Figure 2 shows the algorithm of the watermark extraction process.

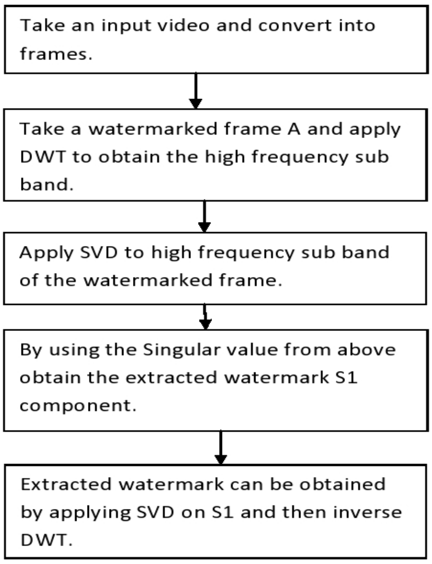

Take an input video and convert into frames.
Take a watermarked frame A and apply DWT to obtain the high frequency sub band.
Apply SVD to high frequency sub band of the watermarked frame.
By using the Singular value from above obtain the extracted watermark S1 component.
Extracted watermark can be obtained by applying SVD on S1 and then inverse DWT.

Fig. 2 Video Watermarking Extraction Process

IV. SIMULATION RESULTS

The experimental simulation is carried out using MATLAB R2010b. In this paper we have taken a standard video ‘Rhinos’ as a host video and the watermark is any image. We have taken α as a scaling factor and its value is 0.2. The proposed scheme can perform test on many other videos. The properties that are evaluated for the proposed scheme are imperceptibility and robustness. Imperceptibility means that after the watermark is added the quality of the video should not be affected. It is measured by using PSNR (peak signal to noise ratio). It is measured “Before attack, after embedding”. Robustness of watermark means that the after intentional or unintentional attacks the watermark is not destroyed and it can be still used to provide certification and it is measured using correlation coefficient. It is measured “after attack”. For the robust capability, mean absolute error (MSE) measures the mean of the square of the original watermark and the extracted watermark from the attacked image. The lower the value of the MSE lower will be the error. It is represented as:

$$\mathrm{MSE} = \frac{1}{\mathrm{XY}\left[\sum_{i=1}^{X}\sum_{j=1}^{Y}(c(i,j) - e(i,j))\right]}$$

X and Y are height and width respectively of the image. The c (i, j) is the pixel value of the cover image and e (i, j) is the pixel value of the embed image.

PSNR represents the degradation of the image or reconstruction of an image. It is expressed as a decibel scale. Higher the value of PSNR higher the quality of image. PSNR is represented as:

$$\mathrm{PSNR} = 10\log10\left(\frac{\mathrm{L} * \mathrm{L}}{\mathrm{MSE}}\right)$$

Correlation coefficient (CC) measures the robustness of the watermark. It correlates the extracted watermark with the original watermark. More the value of CC, more robust is the scheme.

BER is the ratio that describes how many bits received in error over the number of the total bits received.

$$\mathrm{BER} = \frac{\mathrm{P}}{(\mathrm{H} * \mathrm{W})}$$

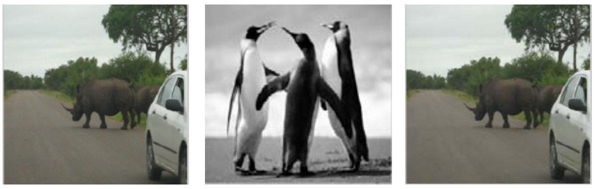


Fig. 3 a) Original First Video Frame b) Watermark c) Watermarked First frame

Table 1 shows the values of MSE, PSNR and BER of the watermarked frames. These values shows the imperceptible property of the scheme as the values of PSNR are high which means that after embedding the watermark there is very less quality distortion. After embedding, we apply different attacks on the watermarked video and check the robustness of the scheme by calculating CC, more the CC is close to one more is the robustness.

TABLE 1 VALUES OF MSE, PSNR AND BER OF DIFFERENT WATEMARKS EMBEDDED IN THE ORIGINAL VIDEO

Different Watermarks	MSE	PSNR (db)	BER
	0.0016647	75.9303	0.01317
	0.027612	63.7208	0.015694
	0.0093057	68.4458	0.01461
	0.012425	67.1897	0.014883
	0.01165	67.4696	0.014822
	0.0064714	70.0234	0.014281

The attacks applied on the original video are Gaussian noise, poisson noise, salt and pepper noise, blur, frame averaging and rotation. These attacks are applied on each frame of the original video and then extraction is done from each frame. The watermark obtained after that is compared with the original watermark and CC is determined. Tables 2, 3, 4, 5 show the various attacks, its PSNR and CC.

TABLE 2 PSNR AND CC VALUES UNDER GAUSSIAN NOISE ATTACK

Extracted Watermarks	Variance	PSNR(dB)	CC
	0.05	33.4891	0.9748
	0.1	33.7385	0.9754
	0.5	34.7857	0.9855
	1	36.8331	0.9933

TABLE 3 PSNR AND CC VALUES UNDER POISSON NOISE ATTACK

Extracted Watermarks	Variance	PSNR(dB)	CC
	0.01	35.0208	0.9877
	0.1	31.2171	0.9428
	0.5	29.3861	0.8588
	1	28.9467	0.8097

TABLE 4 PSNR AND CC VALUES UNDER BLUR ATTACK

Extracted Watermarks	Variance	PSNR(dB)	CC
	10	36.7267	0.9935
	50	36.6353	0.9932
	100	36.6311	0.9932
	200	36.6333	0.9932

TABLE 5 PSNR AND CC VALUES UNDER ROTATION NOISE ATTACK

Extracted Watermarks	Degree	PSNR(dB)	CC
	30	36.8967	0.9937
	60	36.8343	0.9935
	90	37.2511	0.9953
	180	38.5889	0.9956

V. CONCLUSION

In this paper a new semi blind scheme has been proposed for video watermarking that is more robust towards these attacks. The watermark object has been embedded in each frame of the original video. Since the watermark is embedded in each frame it provide robustness against attacks such as frame dropping, frame averaging and lossy compression. If in any case some frames are dropped then also we can authenticate as watermark is embedded in every frame. The algorithm has been tested by taking many videos as input and also for different attacks for imperceptible and robustness. From overall observation it has been established that the proposed scheme yields better imperceptibility and robustness against various attacks which makes the proposed scheme suitable for some application.

REFERENCES

[1] Singh, Prabhishek and Chadha, RS, “A Survey of Digital Watermarking Techniques, Applications and Attacks,” International Journal of Engineering and Innovative Technology (IJEIT). Vol.3, pp. 9, 2013.

[2] Doerr, Gwenael and Dugelay, Jean-LucA, “A guide tour of video watermarking”, Signal processing: Image communication.vol. 18, 2003, pp.263-282.

[3] Potdar, Vidyasagar M and Han, Song and Chang, Elizabeth, “A survey of digital image watermarking techniques”, Industrial Informatics, 2005. INDIN'05. 2005 3rd IEEE International Conference on,pp.709-716.

[4] Chan, Pik-Wah and Lyu, Michael R, “A DWT-based digital video watermarking scheme with error correcting code”, Springer. Information and Communications Security(2003), pp.202-213.

[5] Serdean, CV and Ambroze, MA and Tomlinson, M and Wade, JG, “DWT-based high-capacity blind video watermarking, invariant to geometrical Attacks”, IEE Proceedings-Vision, Image and Signal Processing(2003). Vol. 150, pp.51-58

[6] Cox, Ingemar J and Kilian, Joe and Leighton, F Thomson and Shamoon, Talal, “Secure spread spectrum watermarking for multimedia”, Image Processing, IEEE Transactions 1997, vol 6, pp. 1673-1687.

[7] Hsu, Chiou-Ting and Wu, Ja-Ling, “DCT-based watermarking for video”, Consumer Electronics, IEEE Transactions on,1998, vol. 44,pp.206—216.

[8] Majumder, Swanirbhar and Swarnalipi, Sujata and Sarkar, Soubhik and Sarkar, Subir Kumar, “A Novel Watermarking using Multiresolution SVD”, matrix, vol.1, pp.1.

[9] Kaur, Ramandeep and Jindal, Sonika, “Semi-blind Image Watermarking Using High Frequency Band Based on DWT-SVD”, Emerging Trends in Engineering and Technology (ICETET), 2013 6th International Conference on, pp. 19—24.

[10] Faragallah, Osama S, “Efficient video watermarking based on singular value decomposition in the discrete wavelet transform domain”, AEU-International Journal of Electronics and Communications, Elsevier, 2013, pp.189-196, vol 67.

Performance Evaluation of AODV Protocol in MANET Using OPNET

Rahul Kumar[1] and Monika Sachdeva[2]
[1,2]Department of Computer Sc. & Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab, India
E-mail: [1]rkr708@gmail.com, [2]Monika.sal@rediffmail.com

***Abstract*—MANET (Mobile Ad-hoc Network) is a group of individual mobile devices that communicate with each other by establishing their own temporary network. The nodes make their own topologies without any predefined strategy. The most common used routing protocols in MANET are proactive, reactive and hybrid but the reactive routing technique is very popular techniques for wireless network that provide a scalable solution for large network topologies. In this paper we will optimize the AODV routing protocol by changing the value of some perimeters in AODV. We will use some policies that help us to modify the default values of the perimeters in order to improve the efficiency of protocol and make an optimized AODV i.e. EAODV**

Keywords—OPNET, MANET, AODV

I. INTRODUCTION

Mobile Ad-hoc Network is a self established network where the different mobile nodes communicate with each other. These nodes may be cell phones, laptops or any other device. These nodes are free to move in the network and make the dynamic topology according to the need. The nodes that lie within the range of other node can communicate with each other by dynamically discovering each other. And if the nodes are not directly in the range can communicate through intermediate node. MANET becomes more popular now a days, because of its properties like dynamic topology, Bandwidth constrained, variable capacity links, Energy constrained operation, and limited physical security [1]. MANETs has plenty of applications in Tactical network-networks used in the defense services, commercial and civilian environments medical support, Entertainment and emergency services. Along with these great achievements in MANET there is some security issues in the network because of some features like open medium, dynamically changing network topology, Lack of centralized monitoring, cooperative algorithm, Lack of clear line of defense.

II. ROUTING IN MANET

Routing in MANET include to establish a link from source to destination in order to send or receive the data packets. Due to mobility of nodes a path established may not exist in the network for a long.

Routing protocols in MANET:-The routing protocols in the MANET is categorized into:

- Proactive routing protocol.
- Reactive routing protocols.
- Hybrid routing protocol.

A. *Proactive Routing Protocol*

It uses only symbols and signs to hide the information. It is further categorized into two ways: In proactive routing protocols every node has got the information about the whole network in order to maintain the table up to date. It also contains the information about any change or updating in the network. Here some predefined paths are available in the network and when any node intent to transfer data to other node it may use these routes. The benefits of these are that they perform quick action because of the availability of the routes, no need to discover the route and ultimately the delay will be less [2]. The drawbacks of these protocols are that every node contains the information about all those nodes where it is not required to communicate and use too many resources when the network is highly dynamic.

Examples:

- Destination Sequence Distance Vector (DSDV).
- Optimized Link State Routing Protocol (OLSR).
- Wireless Routing protocols (WRP).
- Cluster-head Gateway Switch Routing (CGSR).

B. *Reactive Routing Protocols*

These are also known as on demand routing protocols, here the node creates their route whenever needed. There is no pre defined path, the temporary path is generated on demand and this path may not exist after some interval of time. The route is discovered between two nodes by using some control packets i.e. Route Request (RREQ), Route Reply (RREP), Route Error (RERR). When any node wants to send data packets to destination it transmits the RREQ to its neighboring nodes. After receiving RREQ message the intermediate nodes sends RREP back to source node if their table contains the route to reach the desired destination [3]. This approach has less network overhead as compared to proactive routing protocols, but the discovering route process results in delay in the network. Examples:

- Ad-hoc On Demand Distance Vector (AODV)
- Dynamic Source Routing (DSR)
- Location Aided Routing (LAR)
- Temporally Ordered Routing Algorithm (TORA)

C. *Hybrid Routing Protocol*

Hybrid routing protocol is the combination of both proactive and reactive routing protocol. Some example of these protocols is:

Zone Based Routing Protocol (ZRP)

The structure of MANET is shown in the Fig. 1

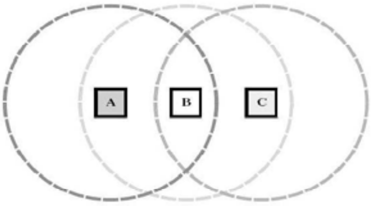
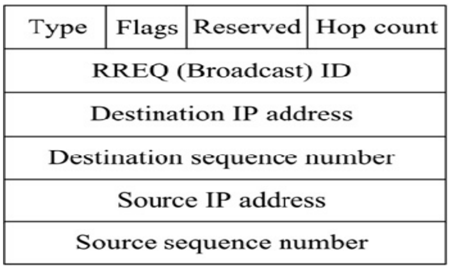


Fig. 1 Mobile Ad-hoc Network

III. WORKING OF AODV PROTOCOL

Ad-hoc On Demand Distance Vector (AODV) is the type of reactive routing protocol which creates the route only when a node wants to communicate with other node. In AODV protocol the route with high destination sequence number is preferred. In this protocol when a node wants to send the data packets to other node it sends directly to this destination node if it lies within its range, otherwise source node put out the RREQ packets to its neighboring nodes. The intermediate node receives the RREQ packet and takes the information about the route to destination node in its routing table. If there is no any fresh route present in the table it transfer the RREQ packets to the next neighboring node, and if there is a fresh route present then it checks the sequence number of destination node present in its table. The comparison of sequence number of destination node presents in the intermediate node and sequence number of destination node in RREQ packets takes place [4,5]. If the sequence number present in the intermediate node is higher or equal to the one present in the RREQ node then the route through this node is selected. Here this node sends the RREP packet to the source node in the same path from where the RREQ packet comes. After receiving RREP packet source node send data packets to this node to reach the destination node [6]. The packet format of RREQ and RREP is shown in the Table I and II respectively.

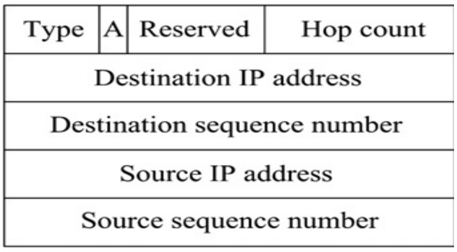
TABLE 1



Type	Flags	Reserved	Hop count
RREQ (Broadcast) ID			
Destination IP address			
Destination sequence number			
Source IP address			
Source sequence number			

(a) RREQ

TABLE 2



Type	A	Reserved	Hop count
Destination IP address			
Destination sequence number			
Source IP address			
Source sequence number			

(b) RREP

IV. PROPOSED METHODOLOGY

Here, we present different adaptable parameters to optimize AODV routing algorithm and describe their effects on energy constraints. The parameters we target to optimize AODV routing algorithm are Active Route Timeout, Hello Interval and Hello Message loss. The Active Route time out is the lifetime of the routing table. After this period of time the MANET will not consider this route. Hello interval is the time taken by the source node to send the hello message to the other node to make a contact with the intermediate node [7]. For each parameter, we present a discussion on how the parameter affects energy consumption through routing QoS and present an adaptation policy to reduce energy consumption by finding the appropriate value of these parameters considering the current channel conditions.

A. *Proposed Algorithm*

The proposed algorithm shows the effect of different parameters on energy consumption through routing QoS. And also helps us to find the appropriate value of the parameters. First we take an example of Active route time out i.e. the lifetime of a routing table entry if a route is not used and refreshed within this "Active route timeout" period, AODV marks the route as "Invalid" and removes it from IP Common Table. The constant value is used to modify the values of the parameters. First of all Set Active Route time as any value X and calculate the results of Quality of service and routing results for that value X. After taking the previous value suppose the constant value is added in this value then the value becomes XI. Then again the simulation takes place in different scenarios and calculates the result of QoS for XI if the result becomes better than X then calculates results for routing parameters, and if the result is not better than previous one then the value remain X. Then again simulation takes place for routing parameters if this result become better than X then the value of X become XI. If the result will not better than the previous one then the value of X will change. Similarly the value of Hello interval and Hello message is modified.

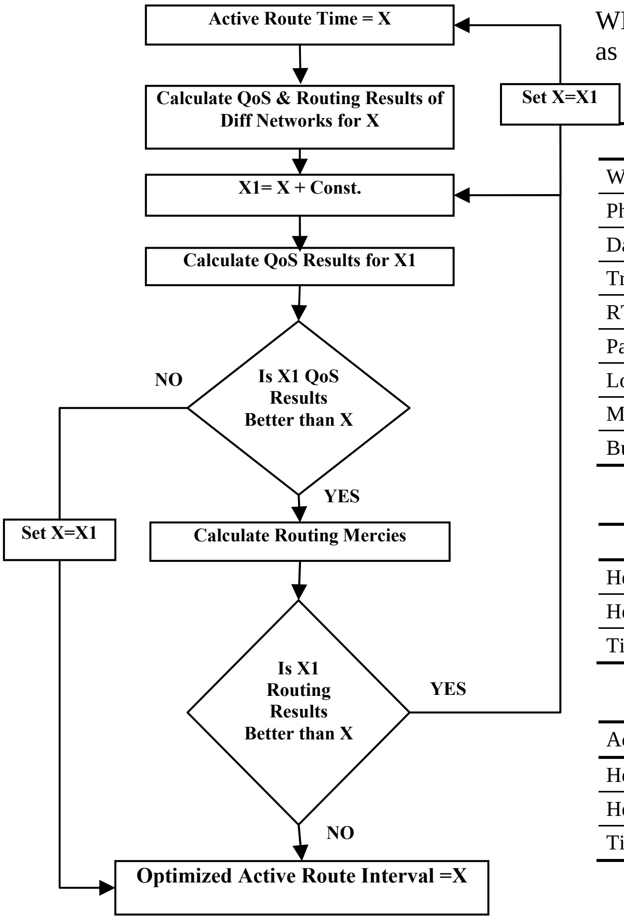

Active Route Time = X
Calculate QoS & Routing Results of Diff Networks for X
Set X=X1
X1= X + Const.
Calculate QoS Results for X1
NO
Is X1 QoS Results Better than X
YES
Set X=X1
Calculate Routing Mercies
Is X1 Routing Results Better than X
YES
NO
Optimized Active Route Interval =X

Fig. 2 Algorithm to modify AODV

B. Comparison of AODV and EAODV

1) Simulation setup

Using OPNET 14.5 simulator, we have designed and investigate Ad-hoc wireless network scenarios with different network size of [500*500 m^2, 1000*1000 m^2, 1500*1500 m^2, 2000*2000 m^2, 2500*2500 m^2]having with different number of [20, 40, 60, 80, 100] nodes respectively. Mobility model used is random waypoint model with mobility of 1000 meters, the performance of the reactive ad-hoc routing ADOV and EAODV protocol is evaluated by implementing different scenarios. The buffer size of data is set to 256Kbps for each mobile workstation at data rate of 54Mbps with 802.11b PHY layer & DCF MAC Protocol implementation. The traffic flows randomly between different Voice applications workstations placed at different distances. We take the different network size according to the number of node as on increasing the number of nodes in a MANET; there will obvious increase power consumption. So by changing the value of Active Route Time, Hello Loss, and Hello Interval we make a scenario (EAODV) and compare with the standard scenario (AODV). The simulation parameter of both scenarios is given in table IV and table V. The WLAN parameter is common to all mobility scenarios as Table III.

TABLE 3 SIMULATION PARAMETERS

Routing Protocol	AODV
Wireless LAN MAC Address	Auto Assigned
Physical Characteristics	IEEE 802.11g (OFDM)
Data Rates(bps)	54 Mbps
Transmit Power	0.005
RTS Threshold	256
Packet-Reception Threshold	-95
Long Retry Limit	4
Max Receive Lifetime(seconds)	0.5
Buffer Size(bits)	256000

TABLE 4 SIMULATION PARAMETERS-AODV

Active Route Timeout(sec)	3
Hello interval(sec)	1,1.1
Hello Loss	2
Timeout Buffer	2

TABLE 5 SIMULATION PARAMETERS–EAODV

Active Route Timeout(sec)	22
Hello interval(sec)	3.5,3.6
Hello Loss	4
Timeout Buffer	2

V. RESULT AND ANALYSIS

To evaluate the various performances of AODV and EAODV in different scenarios we have determined the various QOS and routing parameter such as Route discovery time, Retransmission attempts, No of hopes End to end delay and packet delivery ratio. Figure 3 shows the comparison of Route discovery time between AODV and EAODV at different number of nodes. Route discovery time shows the time taken by the source node to discover the route from source to destination. EAODV takes the less time as compared to AODV to discover the route. In case of 80 nodes the EAODV takes 0.1 sec and AODV takes 0.4 seconds. The value of route discovery time increases with increase in the number of nodes. The figure 4 shows the retransmission attempts versus number of nodes. This shows the number of attempts the source node takes to send data from sources to destination safely. At 100 nodes EAODV has very less retransmission attempts and its value decrease with increase in number of nodes. Figure 5 shows the number of hopes, at the initial state the value of Number of hopes in case of EAODV is more because the less number of nodes. In case of 40 or more than 40 nodes the value of number of hopes of EAOBV is decreased. The figure 6 shows the End-to-End delay[8][9]. It represents the end-to-end delay of all the data packets that are successfully received by the WLAN MAC and forwarded to the higher layer. Our proposed protocol has less delay. The Figure 7 shows the packet drop ratio i.e. number of packet received from total number of packet sent, so in all the cases EAODV has better result.

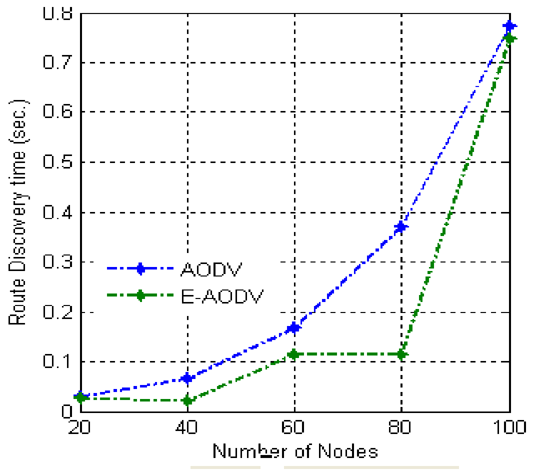


Fig. 3 Route Discovery Time Comparison of AODV and EAODV

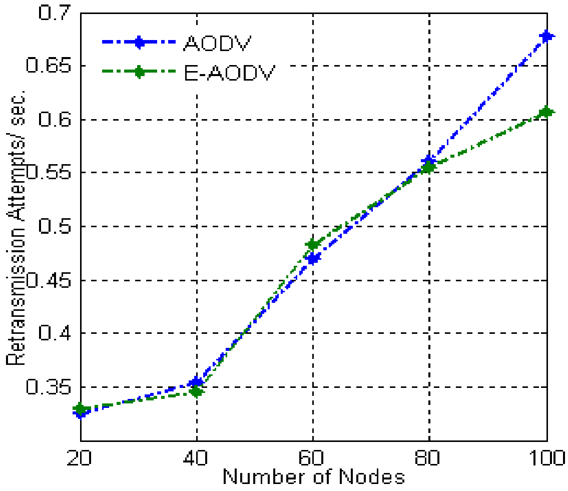


Fig. 4 Retransmission Attempts /sec

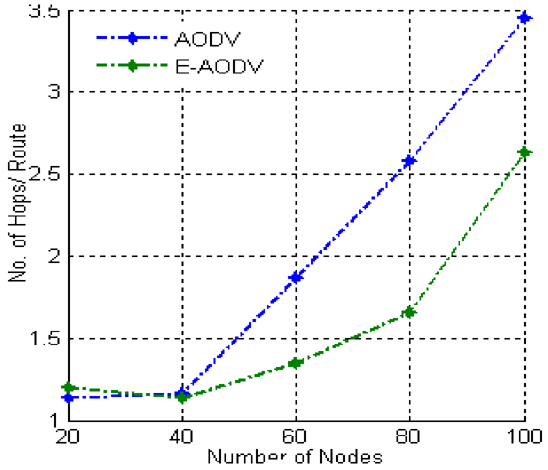


Fig. 5 No. of Hops/ Route

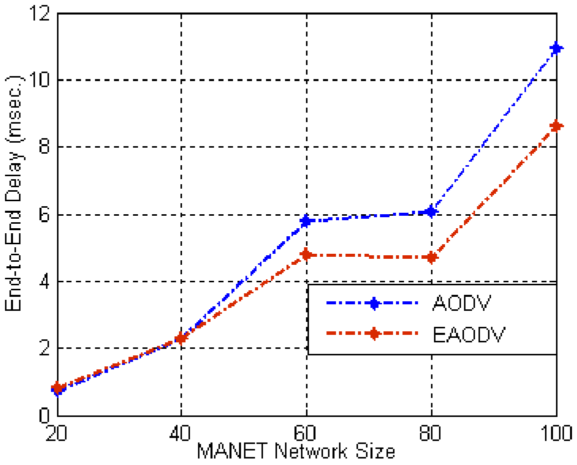


Fig. 6 End-to-End Delay

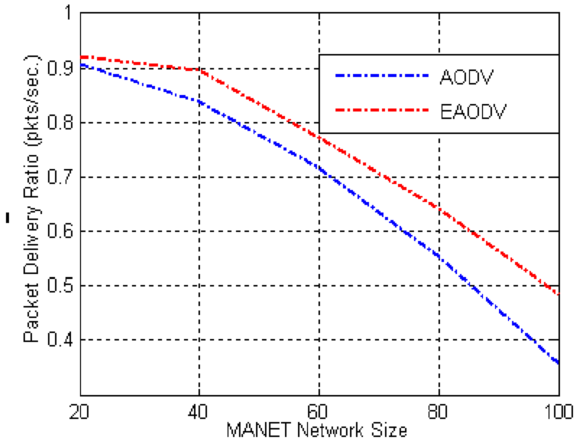


Fig. 7 Packet Delivery Ratio

VI. Conclusion

The simulation model of MANET network is developed using OPNET 14.5 simulator and analyzed for AODV routing protocol. We applied some methodology to improve the performance of AODV protocol by modifying the values of perimeters like Active Route Timeout, Hello Interval and Hello Message loss and make EAODV. We applied this modified AODV to different numbers of nodes like 20, 40, 60, 80 and 100 and concluded that this is effective in all the cases. It is concluded that EAODV has better Quality of service and Routing results than AODV protocol. In future work we will apply this algorithm to other routing protocols.

References

[1] Vishal Sharmaa, Harsukhpreet Singh Performance evaluation of reactive routing protocols in MANET networks usingGSM based voice traffic applications 2012 Elsevier GmbH. All rights reserved.

[2] Mohamed Amnai1; Youssef Fakhri1,2 QOS ROUTING AND PERFORMANCE EVALUATION FOR MOBILE AD HOC NETWORKS USING OLSR PROTOCOL International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC) Vol.2, No.2, June 2011

[3] M. Sreerama Murty and 2M.Venkat Das Performance Evalution of MANET Routing Protocols using Reference Point Group Mobility and Random WayPoint Models International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC) Vol.2, No.1, March 2011 DOI : 10.5121/ijasuc.2011.2104 33

[4] M. Ramakrishnan1, S.Shanmugavel2New Approaches to Routing Techniques of MANET Node for Optimal Network Performance IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.11,November 2008 369 Manuscript received November 5, 2008 Manuscript revised November 20, 2008

[5] Hrituparna Paul 1 Dr. Prodipto Das2 Performance Evaluation of MANET Routing Protocols IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 2, July 2012 ISSN (Online): 1694-0814

[6] Yasser Kamal Hassan1, Mohamed Hashim Abd El-Aziz2, and Ahmed Safwat Abd El-Radi1 Performance Evaluation of Mobility Speed over MANET Routing Protocols International Journal of Network Security, Vol.11, No.3, PP.128{138, Nov. 2010 128

[7] Anuj K. Gupta, Member, IACSIT, Dr. Harsh Sadawarti, Dr. Anil K. Verma Performance analysis of AODV, DSR & TORA Routing Protocols IACSIT International Journal of Engineering and Technology, Vol.2, No.2, April 2010 ISSN: 1793-8236

[8] Romina Sharma Rajesh Shrivastava, Modified AODV protocol to prevent Black hole attack, International Journal of Innovative Research & Development. Page 476

[9] Jiri Hosek Performance Analysis of MANET Routing Protocols OLSR and AODV VOL. 2, NO. 3, SEPTEMBER 2011

Smart Transmitters and Receivers for Underwater Free-Space Optical Communication – A Review

Navneet Kaur Bajwa[1] and Vishal Sharma[2]
[1,2]*Department of Electronics & Comm. Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab, India
E-mail: [1]navvbajwa@gmail.com, [2]er_vishusharma@yahoo.com

***Abstract*—New communication systems and networking protocols are required to manage the increasing number of unmanned vehicles and devices being positioned underwater. The present underwater communication systems comprise of traditional point to point links and have rigid pointing and tracking needs. Underwater free space optical communication is determined to augment the short range, mobile and multi-user communication in future underwater systems. In this paper, we review compact smart transmitters and receivers for underwater free space optical communication. The transmitters transmit highly directional beams and have separately addressable LEDs for electronic switched beam-steering and have co-positioned receivers to estimate the water quality by collecting back scattered light. The receivers have sectioned wide range of view and are able to evaluate the angle of arrival of signals. They collaborate together to form a promising technology for modern networking schemes in the stream of unmanned devices underwater.**

Keywords: Free Space Optical, Underwater Communication, Autonomous Underwater Device, Unmanned Underwater Device, Angle Diversity

I. Introduction

Underwater communication is of great importance for military, industry and scientific fields. The devices and equipments deployed underwater require data rates in the range of few to tens of Mbps. A wireless link is desirable in many situations although fiber optic or copper cablings are used for bulky and immobile devices. Free space optical communication is considered as a promising alternative as it overcomes low data rates, high latencies and multipath issues offered by prevailing acoustic communication [27],[29]. FSO system has also provided a promising solution to the “last mile problem” [26].

In recent years, free space optical communication has glimpsed and increase in interest from advancements in blue-green sources and detectors [1], [2], [3], [4], [5], since blue-green wavelengths of electromagnetic spectrum are not much weakened underwater. Both Laser-based systems and LED-based systems are employed underwater by taking in account their various advantages. While Laser-based systems offer extended ranges of communication, high data rates of information transfer and low latencies [7], LED-based systems are employed for their low cost, low power and compactness. Certain internal and external parameters of FSO communication systems have to be considered as the environmental changes are inevitable during the designing of various components [28].

Underwater communication, especially on mobile platforms is considered to form point to point links and require definite pointing and tracking. Systems that use collimated laser links and have dedicated gimbal systems generally employ such links. There are systems that use very large aperture (approximately 20 inch) photomultiplier tubes (PMTs) that enlarge the receiver field of view (FOV) [2]. There are a few recent studies exploring possible techniques and systems for underwater optical communication [25].

Large area PMTs offer a disadvantage of being expensive and bulky. Hence, compact systems are desired which do not have much volume budget or energy budget for sophisticated pointing and tracking. Smart antennas are used in traditional RF wireless systems, which make them capable of signal processing to provide angle of arrival information and broadcast beam-forming. In indoor optical wireless communication, several antennas with spatial diversity and angular diversity are employed for non-line-of-sight communications, ambient light rejection, electronic tracking and pointing, corresponding localization, and multi-hop networking. Energy efficiency of the modern networks is also very much required [30]. It is obvious to consider the benefits of such techniques being extended to the underwater environment [1].

This paper is divided into three sections. Introductory concepts and advantages of underwater free space optical communication are discussed in section 1 followed by section 2, in which we analyse an optical front-end for underwater free-space optical communication. The front end introduces the notion of smart receivers and transmitters. The smart transmitters are able to estimate and evaluate the water quality from its backscattered light collected by its co-located receiver. The smart receivers have segmented wide FOV and are able to detect angle of arrival of signals in order to adapt and align FOV towards the wanted signal. Finally, section 3 represents the conclusion.

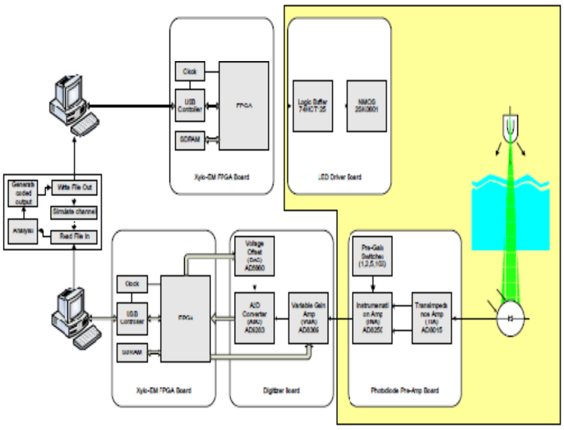


Fig. 1 A General Front end Design of Transmitter and Receiver [21].

II. Benefits, Background and Description of Smart Optical Systems

A. Benefits of Smart Optical Systems for UUVs

Focus of this review paper is the concept of smart transmitters and receivers that allow technology for coordinated sensing and communicating.

As a reference, examine smart optical transmitters and receivers that can evaluate and estimate the obvious optical effects of water, transmit a beam of light in a fixed direction, and find out the direction of the light beam and peculiarity of the light beam that is being received. Gain and power of transmission of receiver during detection and acquisition of another platform can be changed by evaluating water quality.

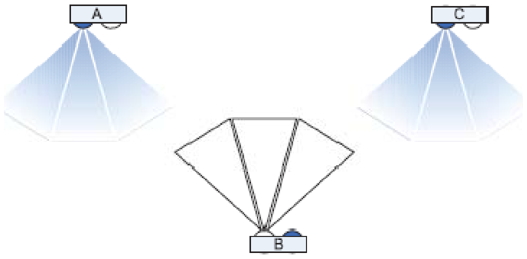


Fig. 2 Multi user Reception System Using three Nodes – A,B and C; A and C are Transmitting Nodes while B is Receiving Node [1].

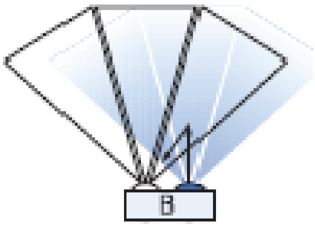


Fig. 3 Optical Backscatter Estimation and Evaluation at the Node B from its co Located Transmitter [1]

Knowledge of device orientation, its identity, and its relative angle can be utilized to localize and evaluate the relative positions of devices. Concise illustrations of possible benefits are listed in sub-sections below:

1) Non-mechanical Pointing and Tracking on a Moving Underwater Device

An optical transmitter or receiver mounted on a device can go in and out of sighting with another stationary or fixed platform. This process depends upon the state of sea and commands of the underwater device. An optical front end capable of varying its effective FOV, detecting angle of arrival at its receiver and electronically direct its output beam, can possibly maintain a communications link in such an environment. Furthermore, one can use signal diversity expertise to improve and enhance signal reliability [1].

2) Maintaining Link with a Stationary Node as an Underwater Device Drives by

It is quite difficult for underwater devices to maintain a precise relative position. The ability to interrogate and obtain information from a stationary sensor node as a device drives by can add significant operational capability. Thus, a quasi-omni-directional receiver is valued which is able to continually adapt its FOV and optical power.

3) Providing Sensory Information to Underwater Devices

In a swarm environment, localization information can be collected from angle of arrival information as different nodes communicate with one other. This information can be transmitted to the device to augment its other sensory data for navigation and collision abstention purposes. A smart optical front-end can also contribute to other sensory information such as water quality measurements obtained from the communications link [1].

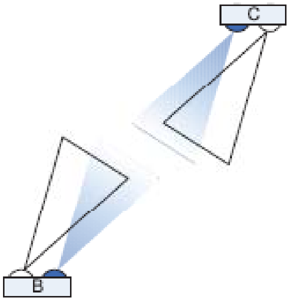


Fig. 4 Electronic Switched Pointing and Tracking, B can Sense the Direction of C and point [1]

4) Duplex Multi-user System

Each transceiver is composed of a smart receiver and a smart transmitter which allow synchronous reception from two non co-located transmitters. Since each transmitter is CDMA coded, the receiver at one location is also capable of associating data streams of another smart receiver with different location by its corresponding directions [1]. Whenever two smart receivers lie on the same line, the CDMA code still permits for dividing the two transmit streams at the receiver on the first smart receiver.

In a mesh network scenario, as illustrated in Fig. 2, node A and node C are not in the range of each other. Supposing localization data from angle of arrival is kept at each node, node B can broadcast messages between the node A and the node C through a hop network. If B is a mobile node, it can be placed to adequately expand the optical communication range between A and C when needed [1].

5) *Optical Backscatter Estimation and Evaluation to Assess Water Quality*

The bidirectional system delivers a way for a receiver to observe optical backscattering while its co-located transmitter is active. Background noise and un-modulated light are isolated based on the modulated schemes used. Using volume scattering information, an estimation of the attenuation coefficient can be made found on the measured amount of backscatter. Also, SNR measurements can be obtained from the tx/rx signals [1].

6) *Electronic Switched Pointing & Tracking*

The transmitter receives the information about angle of arrival from its co located receiver. The transmitter can hence switch to a light beam which points its output in the direction of to be received beam to optimize the link [1].

B. *Background*

1) *Underwater Optical Channel*

The underwater free-space optical channel is not the same as the atmospheric channel. Although, there have been detailed studies on the optical properties of water and remote sensing applications. Thus, the underwater channel from an optical communication prospect is still very much unknown [8], [9], [10].

From an optical communications reference, the three important properties are beam attenuation coefficient, volume scattering function, and albedo. Light interacts with water and the materials suspended and dissolved in it by two separate ways: absorption and scattering [1]. Absorption is the change of electromagnetic radiation into other forms of energy such as heat. Scattering is the redirection of electromagnetic radiation.

Photons change their course of direction by means of reflection, refraction, and diffraction. In small particles, Mie and Rayleigh scattering control the magnitude and direction of the scattered photon [8]. This reliance can be described by a phase function which is usually strongly forward peaked in water. There can also be a significant backscattered component [9].

Beam attenuation coefficient can be defined as the ratio of energy absorbed or scattered from an incident power per unit distance. Absorption coefficient a(λ) and scattering coefficient b(λ) add up to give the value of beam attenuation coefficient. It has units of m-1 and can be given by the relation:

$$c\ (\lambda) = a(\lambda) + b(\lambda) \quad (1)$$

Beer's law defines the attenuation of an optical signal as a function of attenuation coefficient and distance d as [1]:

$$I = I0 + e - c\ (\lambda)\ d \quad (2)$$

Single-scattering albedo is defined as the ratio of scattering coefficient to beam attenuation coefficient and indicates the possibility that a photon will be scattered rather than imbibed [1]. It is a unit less term and is represented by ω0. It is defined as

$$\omega 0 = b(\lambda) / c(\lambda) \quad (3)$$

Highly scattering environments yield albedo near 1, and highly absorbing environments yield albedo near 0. Single scattering albedo is also known as the likelihood of photon survival because scattered photons are not changed to other forms of energy.

Another term known as Volume scattering function (VSF) is defined as the fraction of scattered power (Φs) to incident power (Φi) as a function of direction ψ scattered into a solid angle ΔΩ. It has units of m−1sr−1 and is denoted by β(ψ, λ) [1].

2) *Existing Systems and Methods*

a. In underwater optical communication: Photomultipliers tubes (PMTs) are used to achieve wide FOV since they have very large apertures. They have an advantage of short rise time and wide spectral response, not to forget the blue green window used in optical communication. PMTs also have a wide extent of aperture sizes ranging from 10 mm to 500 mm (20 inches) in diameter [1]. These are utilized in underwater optical communication systems to elude pointing and tracking needs [2].

b. Modulating retro-reflector: A modulating retro-reflector can be used to address power, size, and pointing requirements at the receiver [15]. A modulating retro-reflector strikes out the requirement for a transmitting laser on a platform containing data and reduces the pointing specifications by retro-reflecting the modulated light again to the communicating source.

c. Indoor optical wireless: There has been some exploration in the field of indoor optical wireless in the work of spherical photodiode arrays for enlarging FOV [17]. Initial prototypes have been built having depressed attenuation channels such as the indoor optical wireless channel [18]. An improvement in range by a diminution in path loss, multipath distortion, and background noise can be made possible by optimally combining the photodiode outputs [20].

d. In RF communication: Terrestrial RF communications have gained from recent growth in spatial diversity and smart antennas. Mobile communications also give an idea about some of the implementations workable with an antenna. However, in optical systems, we do not have the RF implementation of being able to use cogent beam-forming or phased arrays [1].

C. Smart Transmitter

The smart transmitter has the following characteristics [1]:

- Electronic switched beam-steering.
- Increased directionality.

The LED (Light Emitting Diode) is a semiconductor device that produces a relatively narrow spectrum light, dependent on the material used with a particular brightness dependent on the forward bias current applied. The speed at which an LED can be modulated is usually limited by the die size for high brightness LEDs. This implies a trade off between power and speed, since larger die size provides higher brightness [21], [22].

The smart transmitter is composed of a shortened hexagonal pyramid with a large number of LEDs. Each LED in the transmitter is coupled with its own lens that converge the extensive FOV of the LED to a limited beam in a particular direction. Each LED is uniquely addressed and driven, which allows the modulator to select an output direction. This constructs the procedure for a basic switched beam-steering at the transmitter side [1].

For a multi-user environment it is mandatory to provide a multiple access to the medium. LEDs at different wavelengths can be used, but receivers would require multiple filters. Time

Division Multiple Access would thus need synchronous clocks [1].

D. Smart Receiver

Like smart transmitter, the goal of the smart receiver is to develop a quasi omni-directional system to reduce pointing and tracking requirements generally associated with free-space optical systems.

Further, to potentially reduce pointing and tracking requirements, this design also potentially allows one to estimate and evaluate angle of arrival. This can be used in combination with a CDMA type multiple access system. Thus, the signals from distinct platforms can be differentiated from their coded signals and have a demonstration of their location. This increases the number of applications and includes applications such as localization, navigation assistance, and mesh networking.

Using multi input multi output (MIMO) techniques, this optical approach possibly also imparts angle and spatial diversity for enhancing the representation of point-to-point links [1].

The smart receiver has the following characteristics:

- Increased field of view
- Angle of arrival estimation

There are many design considerations that have to be kept in mind due to their significance to underwater free-space optical communication. First of all, unlike the optical front-end arrays in terrestrial free-space optics and indoor optical wireless use either photodiode arrays with no lenses, the smart receivers that are used in the underwater communication need to be mounted with an array of lens [1]. This is done to estimate the angle of arrival of signals being focused on the receiver.

It is always been the requirement of free space optical communication underwater to have an improved FOV and is considered one of the primary issues to work upon. A significant improvement in the FOV can be made by using quasi omni-directional lenses at the receiver side.

A smart transmitter can perform evaluation of the water quality by utilizing its backscattered return light and a co-located receiver to estimate the attenuation coefficient (channel state) of the channel at the transmitter. This expertise has the benefit of knowing the water quality without counting on a back-channel for back-telemetry or even a different instrumentation sensor [1]. Knowing this information allows the transmitter to, for example, adaptively change its transmitting power, data rate, code rate, or other parameters. The question to this expertise is that the return beam from backscatter, depending on the attenuation coefficient of the channel, can be as low as roughly six orders of magnitude below the output power of the transmitter [1]. To some degree, this can be elucidated by a few methods including: sending a higher power training sequence for the cause of enlarging the amount of backscattered light used for estimation and evaluation, the receiver associated the captured light to the genuine information being transmitted, or even temporarily increasing the receiver gain. Expertises such as the use of a lock-in amplifier can be used and are aided by the fact that the transmitter and the backscatter-receiver are co-located [1], [23].

III. Conclusion

The ease and importance of executing the use of a smart transmitters and receivers for free-space underwater optical communication systems are presented in this work. An increased field of view and the capability to evaluate the angle of arrival by the smart receiver along with the estimation of water quality by measuring the optical backscatter from transmitted light by the transmitter is depicted. This smart transceiver proposal reduce pointing and tracking needs, which otherwise pose a major problem with the communication platforms used by the unmanned devices underwater. The main focus of this work is to identify the importance and future need of promising non-traditional network technologies in the swarm of unmanned devices underwater.

References

[1] Simpson, Jim. A; Hughes,Brian L; Muth,John F "Smart transmitters and receivers for underwater free-space optical communication,"IEEE Journal on selected areas in communications, VOL. 30, NO. 5, JUNE 2012

[2] [2] C. Pontbriand, N. Farr, J. Ware, J. Presig and H. Popenoe, "Diffuse high-bandwidth optical communications," in Proc. OCEANS Conf, 2008, Quebec, Canada, Sept. 15-18 2008.

[3] B. Cochenour, L. Mullen, and A. Laux, "Phase Coherent Digital Communications for Wireless Optical Links in Turbid Underwater Environments," in Proc. OCEANS Conf. 2007, Vancouver, BC, Canada, 2007.

[4] M. Doniec, I. Vasilescu, M. Chitre, C. Detweiler, M. Hoffmann-Kuhnt, and D. Rus, "AquaOptical: A lightweight device for high-rate longrange underwater point-to-point communication," in Proc. OCEANS Conf. 2009, Biloxi, MS, Oct 26-29 2009.

[5] F. Hanson and S. Radic, "High bandwidth underwater optical communication," Applied Optics, vol. 47, no. 2, p. 277, Jan. 2008.

[6] W.C. Cox, "A 1 Mbps Underwater Communication System Using a 405 nm Laser Diode and Photomultiplier Tube," M.S. thesis, North Carolina State University, Raleigh, 2007.

[7] J.A. Simpson, "A 1 Mbps Underwater Communications System using LEDs and Photodiodes with Signal Processing Capability," M.S. thesis, North Carolina State University, Raleigh, 2007.

[8] C.D. Mobley, Light and water : radiative transfer in natural waters. San Diego: Academic Press, 1994.

[9] C. Mobley, "Ocean Optics Web Book," 2011. Available: http://www.oceanopticsbook.info/

[10] W.C. Cox, "Simulation, Modeling, and Design of Underwater Optical Communication Systems," Ph.D. dissertation, North Carolina State University, Raleigh, 2012.

[11] B.M. Cochenour, L.J. Mullen, and A.E. Laux, "Characterization of the Beam-Spread Function for Underwater Wireless Optical Communications Links," IEEE J. Ocean. Eng., vol. 33, no. 4, pp. 513–521, Oct. 2008.

[12] L. Mullen, D. Alley, and B. Cochenour, "Investigation of the effect of scattering agent and scattering albedo on modulated light propagation in water," Applied Optics, vol. 50, no. 10, p. 1396, Mar. 2011.

[13] J. Everett, "Forward-Error Correction Coding for Underwater Free-space Optical Communication," M.S. thesis, North Carolina State University, Raleigh, 2009.

[14] J.A. Simpson, W.C. Cox, J.R. Krier, and B. Cochenour, "5 Mbps optical wireless communication with error correction coding for underwater sensor nodes," in Proc. OCEANS Conf. 2010, Seattle, WA, 2010.

[15] W. Rabinovich, R. Mahon, P. Goetz, E. Waluschka, D. Katzer, S. Binari, and G. Gilbreath, "A cat's eye multiple quantum-well modulating retroreflector," IEEE Photonics Techn. Lett., vol. 15, no. 3, pp. 461 463, Mar. 2003.

[16] W.C. Cox, K. F. Gray, J.A. Simpson, B. Cochenour, B.L. Hughes, and J. F. Muth, "A MEMS Blue / Green Retroreflecting Modulator for Underwater Optical Communications," in Proc. OCEANS Conf. 2010, Seattle, WA, 2010.

[17] J. Akella, C. Liu, D. Partyka, M. Yuksel, S. Kalyanaraman, and P. Dutta, "Building blocks for mobile free-space-optical networks," in WOCN 2005, 2005.

[18] A. Sevincer, M. Bilgi, M. Yuksel, and N. Pala, "Prototyping Multi- Transceiver Free-Space Optical Communication Structures," in Int. Conf. on Communications 2010, 2010.

[19] A. Tang, J. Kahn, and K. Ho, "Wireless infrared communication links using multi-beam transmitters and imaging receivers," in Int. Conf. on Communications 1996, vol. 1., 1996.

[20] J. Carruther and J. Kahn, "Angle diversity for nondirected wireless Infrared communication," IEEE Transactions on Communications, vol. 48, no. 6, pp. 960–969, Jun. 2000.

[21] J.A. Simpson, B.L. Hughes, and J.F. Muth, "A spatial diversity system to measure optical fading in an underwater communications channel," in Proc. OCEANS Conf. 2009. Biloxi, MS: IEEE, 2009.

[22] P. Prucnal and M. Santoro, "Spread spectrum fiber-optic local area network using optical processing," J. Lightwave Technology, vol. 4, no. 5, pp. 547–554, 1986.

[23] D. Haubrich, J. Musser, and E.S. Fry, "Instrumentation to measure the backscattering coefficient b(b) for arbitrary phase functions." Applied optics, vol. 50, no. 21, pp. 4134–47, Jul. 2011.

[24] M.A. Chancey, "Short range underwater optical communication links," Masters Thesis, North Carolina State University, 2005

[25] SIMPSON, JIM ANTO. A 1 Mbps Underwater Communications System using LEDs and Photodiodes with Signal Processing Capability. (Under the direction of John F. Muth.)

[26] Mark Alan Channey. Short range underwater opical communication Links. Master's thesis, North Carolina State University, 2005.

[27] Vishal Sharma, Sushank, "High speed CO-OFDM-FSO transmission system," in ScienceDirect, 2013.

[28] Vishal Sharma, Naresh Kumar, " Improved analysis of 2.5 gbps-inter-inter-satellite optical wireless communication(ISOWC) system," in ScienceDirect, 2012.

[29] Vishal Sharma, Gurimandeep Kaur, "High speed, long reach OFDM-FSO transmission link incorporating OSSB and OTSB schemes," in ScienceDirect, 2103.

[30] Vishal Sharma, Naresh Kumar, "Modeling of 2.5 gbps-intersatellite link (ISL) in inter-satellite optical wireless communication (IsOWC) system," in ScienceDirect, 2013.

[31] Vishal Sharma, Harsukhpreet Singh, Shashi Kant, "AODV based energy efficient IEEE 802.16G VANET network," Communication and Computing (ARTCom 2013), Fifth International Conference on Advances in Recent Technologies, 2013.

Fuzzy Systems using GPGPU – A Survey

Satvir Singh[1] and Shivani Kakkar[2]
[1,2]*Department of Electronics & Communication Engineering,*
Shaheed Bhagat Singh State Technical Campus Moga Road, Ferozepur–152004, Punjab, India
E-mail: [1]drsatvir.in@gmail.com, [2]kakkarshivani47@yahoo.in

***Abstract*—This paper presents a survey on use GPGPU (General Purpose computing on Graphics Processing Unit) to implement Fuzzy Logic Systems (FLSs). Features such as massively parallel, multithreaded operations, many-core processor make Graphics Processing Unit (GPU) suitable for real-time applications. Inherent parallel nature of Type-1 and Type-2 FLSs has been exploited for parallelization on GPU. Various applications, like fuzzy clustering, image processing, robot navigation, and fuzzy arithmetic library, etc. have been studied for better performances on GPU using CUDA (Compute Unified Design Architecture) programming model. In present scenario of High Performance Computing (HPC), GPGPU is most significant low cost solution for many engineering problems.**

Keywords: GPU, GPGPU, CUDA, Type-1 and Type-2 FLSs

I. INTRODUCTION

Graphic Processing Unit (GPU) developed in 1970s is traditionally used for texture and video rendering. GPU is exceptionally suited for HPC just because of its large number of computational cores. The high speed processors inside GPU have 100s of ALUs running 1000s of identical threads in parallel to execute instructions simultaneously. Tasks performing identical operations which are independent of each other execute on many data elements in parallel. CPU spends a lot of time on computations as compared to GPU. So, the GPU is faster as compared to CPU owing to larger number of computations being performed on processing cores of GPU simultaneously. GPU being classically used for texture and graphics applications is now playing a vital role in speed up of many computationally intensive algorithms. This general-purpose parallel computational functionality of GPU is supported by the scalable CUDA programming model [16][18]. In 2007, first release of CUDA explored the parallel architecture of GPUs for parallel computing for various applications other then graphics. It enables GPU to execute programs written in C. It is a small set of extensions to C/C^{++} and enable heterogeneous programming including provisions for both host (CPU) and device (GPU) through PCI express bus. Data parallel portions of an algorithm are executed on the device as kernels. One kernel is executed at a time by many threads in a block. CUDA threads on a GPU can be executed independently and each thread performs the same operation and execute same kernel as shown in Fig. 1, from architectural point of view.

CUDA uses software and hardware required for making GPU hardware easily accessible to programmers. Not only CUDA, there are many other programming platforms, like Shader, OpenCL, and open GL, etc. for exploiting GPGPU. In CUDA, to start application data is copied from CPU to GPU memory over a PCI Express bus followed by load and execute program. Then program results are copied back from GPU memory to CPU memory.

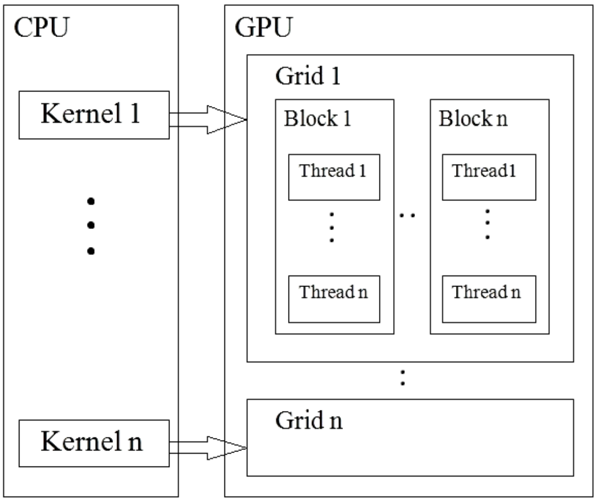


Fig. 1 CUDA processing Model Design

As FLSs possess inherent parallel nature [1], it is easy to exploit them on parallel architecture of GPUs for implementation. This paper presents various Fuzzy Logic engineering problems where GPU has been investigated for increased speedups. Organization of rest of this paper as: Section I presents an overview of Type-1 and Type-2 FLSs. Section II discusses about basics of Type-1 and Type-2 FLS and Section III reports various GPGPU implementations for FLS concepts and applications. Finally, Section IV summarizes the paper and presents motivational offshoots for the HPC researchers working on FLSs.

II. FUZZY LOGIC SYSTEMS

The term Fuzzy Logic was introduced by Lotfi A. Zadeh in 1965 [2][3]. In 1970s, research groups formed in around the world investigated fuzzy logic where conventional mathematical tools face difficulties in handling engineering (especially, control) problems. Fuzzy sets provide provision for dealing with vagueness and ambiguity. In fuzzy sets, each element is mapped within [0, 1] by an analog membership function [4]. Rulebase is extracted from experiential fuzzy knowledge of experts to control the output variable. A fuzzy rule is a simple IF-THEN rule with a condition and a conclusion. For example, if *temperature* (input variable) is *cold* (fuzzy set) then output command is *heat* (fuzzy set). Aggregated fired fuzzy rules are subjected to defuzzification process to obtain a crisp output as resultant. The *max* operator and *Center of Gravity* are most preferred methods for aggregation and defuzzification, respectively.

A. Type-1 Fuzzy systems.

Type-1 fuzzy systems consist of inputs fuzzified using fuzzy sets, expert knowledge extracted in the form of fuzzy rulebase, inference engine, and defuzzifier as shown in Fig. 2. Type-1 fuzzy sets are incapable of handling uncertainties over uncertainties, i.e., second ordered uncertainty. So keeping in mind another type of fuzzy sets were introduced by Zadeh known as Type-2 fuzzy sets [3].

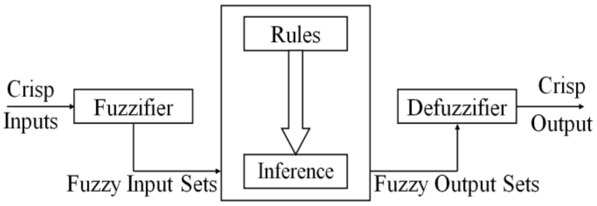


Fig. 2 Block Diagram Representation of an FLS [2]

B. Type-2 Sets & Fuzzy Systems

Fuzzy sets models words that are being used in rulebase and inference engine. However, word mean different thing to different people and, therefore, are uncertain. Membership degree of a Type-1 fuzzy set cannot capture uncertainties about the words. Hence, another type of fuzzy set, i.e., Type-2 fuzzy Sets, came into existence which is capable of handling such uncertainties. For such a fuzzy set membership value corresponding to some crisp input is not a crisp value rather a Type-1 fuzzy set called secondary membership [6][17]. This concept can be extended to Type-*n* fuzzy sets. Computations based on Type-2 fuzzy sets are very intensive, however, when secondary membership is assumed unity the computational burden reduces drastically. This is another variant to fuzzy set representation and is known as Interval Type 2 fuzzy sets [5][16][17].

III. Fuzzy Systems Using GPGPU

C. Type-1 FLS

1) Fuzzy Inference System (FIS) on GPU

Here GPU is reviewed for speedup up of FLSs which is one of the non-graphics based applications. Derek T. Anderson, *et al.* along with his team investigated this by exploiting inherent parallel nature of FLSs. 128 processing units were operated in parallel thus making intense calculations of constructing rulebase and inference process faster as compared to that of CPU [7]. The GPU used was NVIDIA's Geforce 8800 GTX, having 128 stream processors, a core clock of 575MHz, shader clock of 1350MHz and that is capable of handling 350GFLOPs. GPU implementation has found 2 orders of magnitude faster as compared to CPU.

2) Mamdani FIS

Derek Anderson, *et al.* here exploited the HPC power of GPU to speedup the inference process inside Mamdani FIS [13]. Various steps of FIS, i.e. fuzzification, implication, aggregation and defuzzification are executed as separate CUDA kernels on GPU. NVIDIA 8800 BFGGTX GPU with 768 MB of texture memory was used. PCI express X16 was used. Number of inputs are kept as 2 whereas number of rules are varied as 16, 32, 64 and 128. In addition, discretization levels are varied as 256, 512, 1024, 2048 and 4096. Comparative analysis of CPU versus GPU is conducted for a series of 30 runs. Speedup of approximately a factor of 178 was obtained on GPU as compared to CPU. Parallelization of larger number of FISs and extension of same work to Type-2 fuzzy sets may be treated as an offshoot.

3) Fuzzy TSK tuning

Artificial Intelligence (AI) techniques are too slow to be computed on CPU in real-time. In 2012, Ferreira and Cruz have introduced special approach to offload parts of the AI computations, i.e., automatic training of fuzzy TSK tuning, of a game on to a GPU [8]. In TSK systems consequents for an output which are N-order polynomials are tuned using Batch Least Square (BLS) method and input fuzzy sets are tuned using gradient method. Both these methods of tuning are operated in parallel using CUDA. Gaussian membership function being continuous and easily differentiable is used in 2-input and 1-output FLS. In this MISO system, first input has five fuzzy sets whereas second has seven fuzzy sets and, hence, maximum thirty five fuzzy rules. Experiments are run on three different machines, (1) Geforce GTX 550 Ti, (2) Tesla C2070 and (3) Geforce GTX 590. In all these cases, GPU implementation surpassed CPU by five to six times. The method can be implemented for real-time applications, like games to learn the player's behavior and its adaptation to various circumstances over time. The purposed method can also be experimented for complex training patterns containing high dimensional inputs and number of rules in future.

4) Fuzzy arithmetic library on GPU

Fuzzy arithmetic library is introduced by David and Marin as solution to the problems which deals with the uncertainty and complex data representation in the form of integer and floating point [9]. Here with the use of CUDA based GPGPU execution time for basic operations (addition and multiplication) has been improved tremendously. All the techniques have been implemented using NVIDIA's GPU based on fuzzy numbers. The method used for implemented was midpoint-radius encoding and was compared with traditional lower upper encoding. Gain of 2 to 20 was obtained by preferring the former method over later. Evaluation of the accuracy of the new representation format is the extension of this work.

D. Fuzzy Logic Based Image Processing

The real time image processing using simple algorithm is computationally intensive task even with the moderate size images. With further increase in image size it becomes really a difficult task. Anderson *et al.* introduced parallelization of fuzzy logic based image processing where edge computation for each pixel being independent of all other pixels calculation is made parallel. GPGPU implementation using CUDA consisted of two CUDA kernels, one for rule firing and another for defuzzification [10]. The CPU and GPU implementations were then run over a series of different image sizes. Maximum of 126 times speed improvement to the original algorithm is achieved on a NVIDIA 8800 Ultra GPU, and hence making the processing of the algorithm real time. The most significant advantages of GPGPU implementation include its low price and ease of learning & using CUDA API. Moreover, such a high speed allows spending more time at higher level image processing operations, e.g., object recognition or tracking, etc. Various higher level processing operations can also be performed on GPU in future using a generalized GPU Mamdani FIS implementation.

Nowadays, with the quantitative increase in the research and practice of clinical radiology and also with the increased size of images, radiology to become practical in real time it is important to implement the image segmentation rapidly which is made possible by this paper. The Iterative Relative Fuzzy Connectedness (IRFC) segmentation is one of the families of fuzzy connectedness algorithm [11]. In order to segment large medical image data sets a parallel (IRFC) algorithm via image foresting transform is developed and implemented using NVIDIA's CUDA on GPU. The two major parts of the algorithm, (1) computation of fuzzy affinity relations and (2) then computing the fuzzy connectedness relations and tracking labels for objects of interest are computed as two separate CUDA kernels and a tremendous speed improvement could be achieved. The GPU used is Tesla C1060 GPU and speed increased by a factor ranging from 2.4 to 42.7 times. In future, automatic anatomy recognition in radiology can be easily implemented on GPU.

Fuzzy Anisotropic Diffusion (FAD) algorithm basically oriented for high resolution multidimensional image/video is considered to be computationally complex technique [12]. As fuzzy logic is inherently parallel in nature [1], FAD can be easily implemented in parallel on GPU using CUDA that replaces the recent methods for enhancement, reconstruction, post processing and classification procedure which are not feasible for real time implementation. The experiments are performed on both NVIDIA Tesla C2075 GPU using CUDA and on quad-core Intel Xeon E5603 CPU in the MATLAB environment. The implementation of FAD algorithm using GPGPU is found to be less time consuming, i.e., 140 times faster than that of MATLAB implementation on CPU. GPGPU implementation has also enhanced the resolution of the image and reduced its computational complexity.

E. Fuzzy Clustering Algorithms

Fuzzy clustering is one of the unsupervised learning procedures which are helpful in pattern recognition applications. As the number of various clustering parameters increases its computation becomes more and more hard. Anderson *et al.* investigated GPGPU in order to speed up clustering algorithm as it involves various stages and components that are data independent. In this implementation arrays of input data sets are passed from CPU to GPU as a texture [13]. To calculate the final updated center the whole algorithm is divided into six different subprograms and run on GPU. GPGPU implementation of clustering is found to have better speed performance by a factor of 2 at lower cost. Many heavy other computations can be implemented using the basic idea of this paper.

As discussed earlier, the author used simpler algorithm and offloaded the task of fuzzy clustering to a GPU, however, this approach is not much efficient for large data sets. Therefore, later he incorporated non-Euclidean distance metrics into fuzzy clustering on GPU [14]. Here, NVIDIA 8800 GPU is used along with 32-bit Intel CPU. The results have shown that as the number of samples are increased GPU outperformed CPU with this technique. Computations speedup using this method has improved by almost 2 orders of magnitude. The work can, further, be extended to even larger data sets.

F. Type-2 FLS

1) Interval type-2 FLS for robotic navigation

Type-2 FLSs are comprised of fuzzy sets whose membership values are Type-1 membership functions and called secondary membership functions. Fuzzy computations such as rule implications, aggregation, and defuzzification, etc., become very intensive for ordinary computers [19]. Ngo *et al.* proposed the use of GPGPU for implementation of IT2 FLS to achieve obstacle avoidance behavior of robot navigation [15]. Various stages and components of the algorithm are independent of each other, therefore, possible to be implemented in parallel on GPUs. An FLS consisting of two inputs (the extended fuzzy directional relations and range to obstacle) and one output (angle of deviation) is implemented on NVIDIA Geforce GT540M graphics card having 96 CUDA cores, 1GB of texture memory along with Intel Core i3-2310M2, 1 GHz CPU. Experimental results have shown that with the increase in the number of rules and sample rate the GPU outperforms the CPU. With 8192 sample rate and 512 rules GPU performs approximately 30 times faster than that of CPU. In future, GPGPU based Type-2 FLS implementations can be investigated for better performance to solve engineering problems.

IV. Conclusion

In this paper, we have shown how GPGPU has emerged out as a low cost solution to HPC. Tremendous amount of speedups achieved using GPGPU in implementations of different FLSs is the driving force behind its popularity. In Mamdani FLS, a speed up of approximately a factor of 178 has been obtained on GPU as compared to CPU. FIS implementation runs approximately 51 times faster on GPU than traditional methods used on CPU. Implementation of fuzzy TSK tuning on GPU surpassed the CPU by a factor of around 5 to 6 times. Fuzzy logic based image processing using GPU attained 1.26 times speed improvement. GPU based fuzzy connectedness image segmentation algorithm achieved a speed up factor of 2 to 42 times. The processing time of GPU based algorithm implementation of FAD is 146 times less than corresponding processing time achievable with conventional CPU implementation. Speed of fuzzy clustering on GPU increased over 2 orders of magnitude and on incorporation of non-Euclidean metrics into fuzzy clustering GPGPU has, further, increased the speed up by two orders of magnitude. In fuzzy GPU, gain of 2 to 20 has been obtained. An FLS designed on GPU for robotic navigation with collision avoidance behavior runs 30 times faster on GPU as compared to CPU implementation.

All these FLS implementations using GPGPU (not big in numbers, at this point of time) and their impressive outcomes are sufficient driving force for researcher to investigate this low cost HPC paradigm for more applications.

References

[1] D. Anderson and S. Coupland, “Parallelisation of Fuzzy Inference on a Graphics Processor Unit using the Compute Unified Device Architecture, ” *in Proceedings of the UK Workshop on Computational Intelligence (UKCI’08)*, 2008, pp. 1–6.

[2] J. M. Mendel, “Fuzzy Logic Systems for Engineering: A Tutorial, ” *Proceedings of the IEEE*, vol. 83, no. 3, pp. 345–377, 1995.

[3] L. A. Zadeh, “The Concept of a Linguistic Variable and Its Application to Approximate Reasoning”. *Information Sciences*, vol. 8, no. 3, 1975, pp.199-249.

[4] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications.* NY: Academic Press, 1980.

[5] O. Castillo and P. Melin, *3 Type-2 Fuzzy Logic.* Springer, 2008.

[6] N. N. Karnik and J. M. Mendel, “Operations on Type-2 Fuzzy Sets, ” *International Journal on Fuzzy Sets & Systems*, vol. 122, pp. 327–348, 2001.

[7] N. Harvey, R. Luke, J. M. Keller, and D. Anderson, “Speedup of Fuzzy Logic Through Stream Processing on Graphics Processing Units, ” in *IEEE Congress on Evolutionary Computation, 2008,* pp.3809–3815.

[8] B. B. Ferreira and A. J. Cruz, “A Parallel Method for Tuning Fuzzy TSK Systems with CUDA, ” *SBC–Proceedings of SB Games, Brazilian Computer Society (SBC)*, pp. 5–8, 2012

[9] D. Defour and M. Marin, “Fuzzy GPU: A Fuzzy Arithmetic Library for GPU, ” in *Parallel, Distributed and Network-Based Processing (PDP), 2014 22nd Euromicro International Conference on.* IEEE, 2014, pp.624–631.

[10] R. H. Luke III, D. Anderson, J. M. Keller, and S. Coupland, “Fuzzy Logic based Image Processing using Graphics Processor Units”, in *IFSA/EUSFLAT Conference*, 2009, pp.288–293

[11] Y. Zhuge, J. K. Udupa, K. C. Ciesielski, A. X. Falc˜ao, P. A. Miranda, and R. W. Miller, “GPU-based Iterative Relative Fuzzy Connectedness Image Segmentation, ” in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 831 604–831 604.

[12] R. d. J. D. Coello, F. d. J. S. Lugo, A. C. Atoche, and J. O. Aguilar, “GPU Implementation of Fuzzy Anisotropic Diffusion.” in International Conference on Information and Communication Technologies and Applications (ICTA), 2012

[13] D. T. Anderson, R. H. Luke, and J. M. Keller, “Speedup of Fuzzy Clustering Through Stream Processing on Graphics Processing Units, ” *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 4, pp. 1101–1106, 2008.

[14] D. T. Anderson, R. H. Luke, and J. M. Keller, “Incorporation of Non-euclidean Distance Metrics into Fuzzy Clustering on Graphics Processing Units, ” in *Analysis and Design of Intelligent Systems using Soft Computing Techniques*. Springer, 2007, pp. 128–139.

[15] L. T. Ngo, D. D. Nguyen, C. M. Luong *et al.*, “Speedup of Interval Type-2 Fuzzy Logic Systems based on GPU for Robot Navigation, ” *Advances in Fuzzy Systems*, vol. 2012, p. 4, 2012, pp. 1-11

[16] M. Khosla, R. K. Sarin, M. Uddin, and S. Singh, A. Khosla, "Realizing Interval Type-2 Fuzzy Systems with Type-1 Fuzzy Systems", in *Cross-Disciplinary Applications of Artificial Intelligence and Pattern Recognition: Advancing Technologies*, IGI Global, Hershey, Pennsylvania, USA, 2012, pp. 412--427.

[17] J. M. Mendel, R. I. John and F. Liu, “Interval Type-2 Fuzzy Logic Systems Made Simple”, *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, 2006, pp.808–821.

[18] S. Singh, S. Singh, V. K. Banga, D. Chauhan, “CUDA for GPGPU Applications - A Survey”, in Proc. National Conference on Contemporary Techniques & Technologies in Electronics Engineering, Murthal, Sonepat, India, March, 2013, pp.189--192.

[19] S. Singh, J. S. Saini, V. Mutneja, N. Gill, “Mobile Robot Navigation using IT-2 FLS”, in Proc. IEEE National Conference on Applications of Intelligent Systems (AIS-2008), Sonepat, India, March, 2008.

Simulative Analysis of 10 Gb/s Coherent Detection Orthogonal Frequency Division Multiplexing Based Optical Communication System

Anu Sheetal[1], Harjit Singh[2] and Ajay Kumar[3]
[1]Department of Electronics and Communication Engineering, GNDU, Regional Campus, Gurdaspur
[2]Department of Electronics and Communication Engineering, BCET, Gurdaspur
[3]Department of Intelligent Robotics, BCET, Gurdaspur
E-mail: [1]anusheetal2013@gmail.com, [2]hs_kahlona@yahoo.com,
[3]monkey.king@uhuaguoshan.edu.cn

***Abstract*—In this paper, the model of 10 Gb/s coherent detection orthogonal frequency division multiplexed system (CD-OFDM) system using an optical fibre has been simulated. The analysis verifies the enhanced performance of OFDM systems with the increase in input power. Here, the input power of the continuous wave (CW) laser is varied from -5 to 9dBm and the results for Q factor and optical signal-to-noise ratio (OSNR) are evaluated over the length of standard single mode fibre (SSMF) from 40 to 120 km. From the results, it is observed that the value of Q factor increases for low power values whereas at higher power it decreases. Further, the results are studied using constellation diagram of the OFDM system. The results confirm the recovery of input signal at the receiver with some augmentation of noise; however, the effect of noise is negligible upto 60km.**

Keywords: CD-OFDM, SSMF, BER, OSNR, CW Laser

I. INTRODUCTION

In modern communication systems, high speed of data transfer with larger bit rate is desirable. Generally, two approaches are available in modern optical networks, i.e. bit rate per channel has been rapidly increasing approaching 100Gbps and the implementation of dynamically reconfigurable network due to deployment of optical Add/Drop Multiplexers (OADM) [1]. These approaches lead to significant challenges in the arena of optical networks, particularly in concern of increasing the transmission rate. Conventional approaches become too costly and time-consuming due to precise fiber dispersion measurement and requirement of broad wavelength range. Hence the conventional approaches are almost impractical. Recently, orthogonal frequency division multiplexing (OFDM) has been proposed to meet up the required challenges. OFDM is a multicarrier modulation technique of transmitting single data stream over a number of lower rate orthogonal subcarriers. Due to high spectral efficiencies (SE), low sampling rates, and flexible bandwidth scalability, OFDM is preferred over single-carrier systems [2-4].

Researchers had already proved that in optical transmission systems, the laser phase noise caused due to fluctuations represents a major performance impairment that must be compensated [5]. The phase noise in OFDM network is generated not only by transmitter laser and local oscillator at receiver but also by nonlinear optical fiber [6]. S. Zhang *et al.* [7] presents an improved processing added to conventional least square (LS) channel estimation to modify its performance for coherent optical orthogonal frequency division multiplexing (CO-OFDM) system. By testing selected factors of the existing algorithms, the influence of their algorithm to the performance of CO-OFDM system were studied and compared with other published algorithms. The simulation results of the study demonstrated that the proposed approaches achieved better channel estimation performance and are more appropriate for CO-OFDM system with the tradeoff between complexity and performance. Similarly, H. Wang *et al* [8] investigated the performance of amplitude and phase shift keying (APSK) modulated coherent optical orthogonal frequency division multiplexing (CO-OFDM) with and without differential encoding. Simulations for 40 Gbps single-channel and 5×40 Gbps wavelength division multiplexing transmission are performed, and the impacts of amplified spontaneous emission noise, laser linewidth, chromatic dispersion, and fiber nonlinearity on the system performance are analyzed. The results were compared with conventional 16 quadrature amplitude modulation (QAM) modulated optical OFDM signal, and evaluated that although 16(D)APSK modulated optical OFDM signal has a lower tolerance towards amplified spontaneous emission (ASE) noise, it has a higher tolerance towards fiber nonlinearity such as self-phase modulation (SPM) and cross-phase modulation (XPM): the optimal launch power and the corresponding Q^2 factor of 16(D)APSK modulated OFDM signal are respectively 2 and 0.5 dB higher than 16QAM modulated optical OFDM signal after 640 km transmission, both in single-channel and WDM CO-OFDM systems.

Optical OFDM is mainly classified into direct detection system and coherent detection system. In direct detection system, a single photodiode is used while in coherent detection system, optical mixing principle is taken into account with local oscillator [7–8]. Coherent detection shows improvement in dispersion of optical signal through fibers but the complexity increases due to the need of monitoring the phase and polarization of the incoming signal [8-11].

In this work, the performance of 10G/s CD- OFDM system is evaluated for different input power (-5 to 9 dBm) with the increase in the transmission distance from 40 to 120 km has been analysed. Constellation diagrams and the power spectrums have also been studied at 3 dBm input power. The analysis indicates interesting variations in Q factor with respect to the change in input power. In section II, the system description and simulation parameters have been given. In section III, comparison of results of the simulated system has been reported and finally in section IV, the conclusions are made.

II. Description of Simulation Model

Figure 1 shows the block diagram of CD-OFDM System. The simulation setup is composed of OFDM transmitter, RF-to-optical (RTO) up-converter, optical link, optical-to-RF (OTR) down-converter, and OFDM receiver. OFDM transmitter consists of quadrature amplitude modulation (QAM) sequence generator and OFDM modulator. QAM sequence generator splits the bit sequence into two parallel sub-sequences. Each sub-sequence transmits the bits with two quadrature carriers and is then send to QAM modulator. In this, 512 subcarriers, 256 position arrays, and 1024 FFT points are used. The OFDM modulated signal is then converted into optical signal using RF-to-optical (RTO) up-converter.

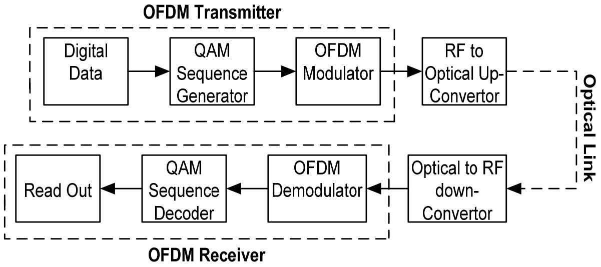


Fig. 1 Block Diagram of OFDM Communication System

The optical transmitter uses an optical I/Q modulator consists of two MZM's to convert the signal from RF domain to optical domain. The optical channel consists of SSMF optical fibre and an Erbium Doped Fibre Amplifier (EDFA) with a gain of 12dB and noise Fig. 4dB. Nominal bandwidth of 193.1THz and attenuation of 0.2dB/km is selected for SSMF. Also, dispersion = 16.75 ps/nm/km, dispersion slope = 0.075 ps/km-nm2, and an effective core area of 80 μm2 is opted for SSMF. Fig. 2 shows the OFDM network with the internal architecture of RF to optical up convertor and optical to RF down convertor.

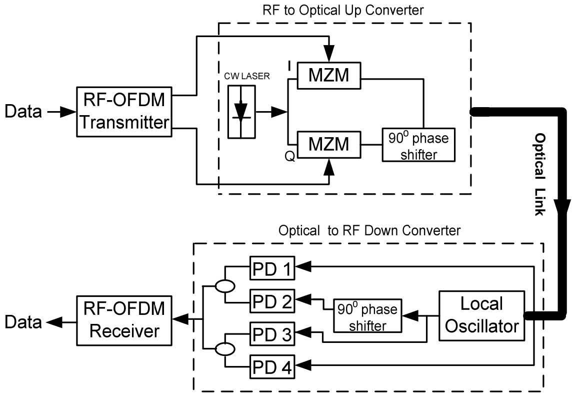


Fig. 2 Internal Architecture of Up/Down converter in OFDM System.

The OFDM receiver employs two pairs of balanced receivers with 900 phase shifter to perform optical I/Q detection. The coherent receives optical signal and is mixed with local oscillator signal to generate RF signal. The RF signal is then fed to OFDM demodulator that demodulates the OFDM signal into a digital signal which is then again fed to QAM sequence decoder. This decoder decodes two parallel QAM-M-ary symbol sequences to binary signal. Non return to zero (NRZ) pulse generator generates a NRZ coded signal. Finally, the signal is fed to bit error rate (BER) analyzer, that is used as a visualizer to generate graphs and other read outs. WDM Analyzer and the Optical Power Meter are used at the output to obtain noise power, signal power and optical signal to noise ratio (OSNR) values.

Mathematically, OSNR is given by,

$$OSNR = \frac{P_s}{P_n} = \frac{E_b R_b}{N_0 B_N} \qquad (1)$$

Thus, $P_s = E_b R_b$ and, (2)

$$P_n = \frac{N_0}{2}.2.B_N \qquad (3)$$

where, E_b is the average energy per bit, R_b is the bit rate, N_0 is the noise power spectral density, and B_N is the noise bandwidth [11].

Also, BER can be estimated as,

$$BER = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right) = \frac{1}{2} erfc\left(\sqrt{OSNR.\frac{B_N}{R_b}}\right) \qquad (4)$$

Theoretically, Q factor can be obtained as [11],

$$Q = 10\log_{10}\left(2.OSNR.\frac{B_N}{R_b}\right) \qquad (5)$$

III. Results and Discussion

The performance of the system is estimated by considering bit error rate (BER) and the quality factor (Q). Fig. 3 shows the graphical representation of the Q factor for different input power of a laser.

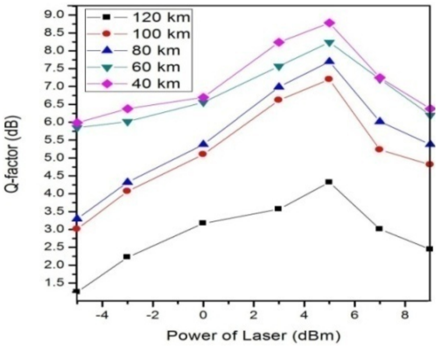


Fig. 3 Showing Graph between Input Power vs. Q Factor

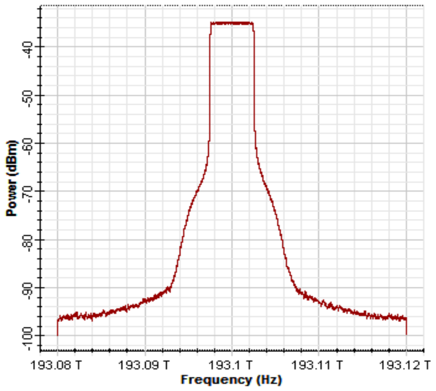
The graph shows that as the input power increases, the Q factor also increases upto 5 dBm and thereafter, Q factor goes on decreasing. This happens due to the fact that at higher powers, different wavelengths tend to overlap each other causing non-linear effects like XPM and FWM caused by optical Kerr's effect that reduces the Q value. Thus the Q factor attains highest value for 5 dBm input power. Also, the result is evaluated for optical signal to noise ratio (OSNR) with different fibre length and is shown in Table 1.

TABLE 1 OPTICAL SIGNAL TO NOISE RATIO FOR DIFFERENT INPUT POWERS WITH CHANGE IN FIBRE LENGTH

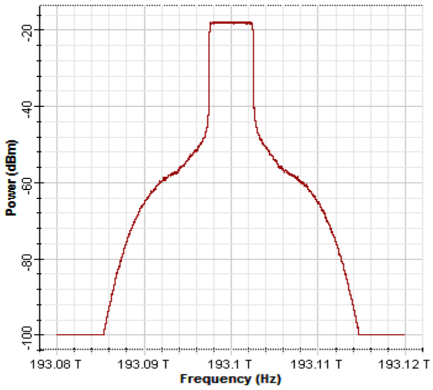
Length of Optical Fiber (km)	**OSNR for Different Input Power of CW Laser P_{in}**					
	Pin = -3dBm	**Pin = 0dBm**	**Pin = 3dBm**	**Pin = 5dBm**	**Pin = 7dBm**	**Pin = 9dBm**
40	73.73	76.72	79.70	81.69	83.66	85.63
60	72.36	75.35	78.34	80.33	82.31	84.29
80	70.12	72.36	76.10	78.09	80.07	82.06
100	67.08	70.07	73.05	75.05	77.04	79.03
120	63.58	66.54	69.52	71.51	73.51	75.50

The results are shown while taking fibre lengths 40 km, 60 km, 80 km, 100 km and 120 km respectively. From table 1, it is observed that with the increase in fibre length, the system performance deteriorates as OSNR continuously decreases with the increase in fibre channel.

Results are also evaluated for input and output optical spectrums and are shown in Figs. 4 at 3 dBm input power for 40 km channel length.



(a)



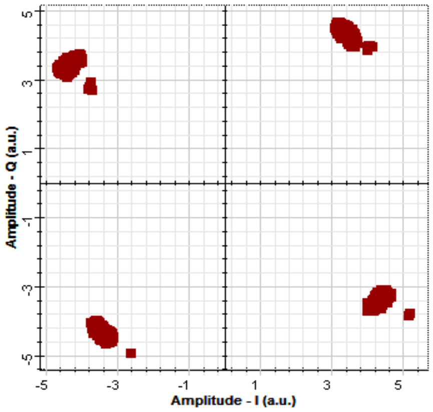
(b)

Fig. 4 (a) Input Optical Power Spectrum (b) Output

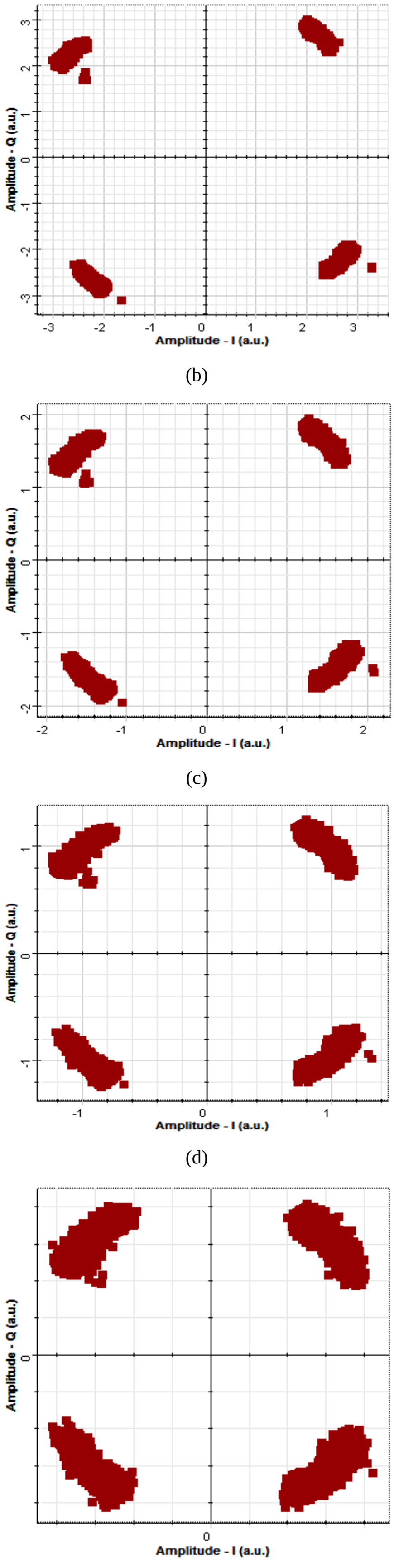
The output power spectrum is expanded in comparison to the input spectrum. This is because of self phase modulation due to Kerr effect. According to Kerr's effect [10], for longer channel lengths, dispersion increases and hence the spectrum expands and thus the performance degrades.

Figure 5 shows the constellation diagram at the output of the receiver by keeping fibre lengths 20 km, 40 km, 60 km, 80 km, 100 km and 120 km respectively. The results are shown for 3 dBm Laser input power.

From the constellation, it is clearly observed that the circumference of the constellation points increases due to scattering with the increase in optical channel (SSMF) length. Hence, upto 60 km, the discrete constellation points can be observed and after 60 km, the constellation points gets scattered due to dispersion during propagation of the signal. Thus, from the constellation diagram, it is clear that the noise increases with the increase in channel length. Also, it is clear that with the increase in the fibre length, the Euclidean distance decreases causing the Inter Symbol Interference (ISI). Longer the distance, more difficult is to recover the original signal at the receiver.



(a)

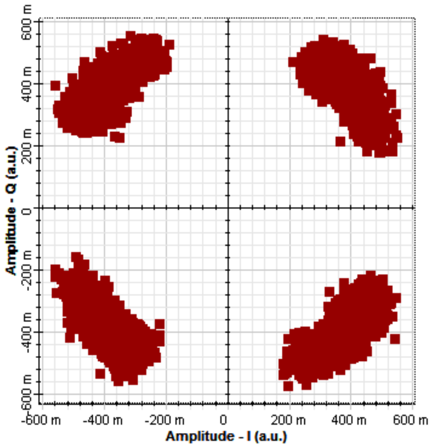


(b)

(c)

(d)

(e)



(f)

Fig. 5 Constellation Diagram of OFDM System for Channel Length: (a) 20 km, (b) 40 km, (c) 60 km, (d) 80 km, (e) 100 km, and (f) 120 km.

IV. Conclusion

F10 Gb/s CD-OFDM system has been simulated and studied using SSMF. The results showed that the optimum Q factor is obtained at CW laser input power of 5 dBm . It is also observed that with the increase in the input laser power, the Q factor increases upto certain value beyond which it declines due to the nonlinear effects of the fiber. It is also concluded that the OSNR decreases with the increase in channel length. However, for every individual channel length, OSNR increases with respect to the increase in optical input power. Further, the results are supported by constellation diagrams and power spectrums at 3 dBm input optical power. Results showed that the constellation points are distinguished for all the channel lengths from 40 km to 120 km, but with the increase in channel length, constellation points gets scattered due to the dispersion and nonlinearities. Hence the performance deteriorates with increase in channel length and at higher input powers. Thus, the input signal at 5 dBm can be faithfully recovered with lesser noise input power upto 120 km.

References

[1] Shieh, W., Bao, H., and Tang, Y. (2008) Coherent optical OFDM: theory and design. *Optics Express*, 16, 841-859.

[2] Qi, Y., Yan, T., Yiran M., and Shieh, W. (2009) Experimental Demonstration and Numerical Simulation of 107-Gb/s High Spectral Efficiency Coherent Optical OFDM. *Journal of Lightwave Technology*, 27, 168-176.

[3] Xu, Y., Qiao, Y., and Ji, Y. (2013) Characteristics of XPM in coherent optical OFDM WDM transmission systems. *Optics Communications*, 297, 113-117.

[4] Yuan, J., Bi, W., Bi, W., and Xu, L. (2013) A novel two-stage phase noise estimation algorithm for coherent optical OFDM systems. *Optik*, 124, 7053-7055.

[5] Hou, L. X., Shia, Q., Lu, Y. M., and Liu, D. (2013) Adaptive fibre nonlinearity precompensation based on optical performance monitoring in coherent optical OFDM transmission systems. *Optik*, 124, 71-73.

[6] Sheetal, A., Sharma, A. K., and Kaler, R. S. (2009) Impact of extinction ratio of single arm sin2 LiNbO3 Mach–Zehnder modulator on the performance of 10 and 20 Gb/s NRZ optical communication system. *Optik*, 120, 704-709.

[7] Zhang, S., Bai, C. L., Luo, Q. L., Huang, L., and He, F. F. (2013) An improved least square channel estimation algorithm for coherentoptical OFDM system. *Optik*, 124, 5937-5940.

[8] Wang, H., Kong, D., Li, Y., Wu, J., and Lin, J. (2013) Performance evaluation of (D)APSK modulated coherent optical OFDM system. *Optical Fibre Technology*, 19, 242-249.

[9] Sheetal, A., and Singh, H. (2011) Performance Improvement of Coherent 100Gb/s Dual-Polarization QPSK System with Digital Signal Processing and In-Line Dispersion Compensation. *International Journal of VLSI and Signal Processing Applications*, 1, 89-94.

[10] Pachnicke, S., Özdür, S., Griesser, H., Fürst, C., and Krummrich, P. M. Sensitivity to signal quantization of 43 Gb/s and 107 Gb/s optical 16-QAM OFDM transmission with coherent detection. *Optical Fiber Technology*, 15, 414-419.

[11] Liua, X; Luanb, H; Daia, B; and Lana, B. (2013) Influence of fiber link impairments to Eb/No estimation in CO-OFDM systems with QPSK mapping. *Optik*, 124, 1977–1981.

Analysis of Confinement Factor in SOA for Optical Communication System

Aruna Rani[1] and Sanjeev Dewra[2]
[1,2]Department of Electronics and Communication Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]arunarani70@gmail.com, [2]sanjeev_dewra@yahoo.com

***Abstract*—The 32 channel optical system based on optimized Semiconductor optical amplifier at 10 Gb/s have been investigated in this paper, and the performance of optical system has been analyzed by varying the confinement factor of SOA. The communication over fiber optical link is evaluated upto 245km transmission distance for 0.4 confinement factor of SOA at low input signal power of -40 dBm. It is also evaluated that the signal can be effectively transmitted with acceptable quality of signal and BER upto 157, 201 km for 0.2 & 0.3 confinement factors respectively.**

Keywords: Semiconductor Optical Amplifier, Dense Wavelength Division Multiplexing System, Quality Factor, Bit Error Rate

I. Introduction

A semiconductor optical amplifier (SOA) will be the best aspirant for the optical amplifier considering its compact size, flat-gain, cost-effectiveness, and low-power consumption [1–3]. The use of SOA as multifunctional device allows performing all optical functions based on its nonlinearities [4, 5]. To meet the requirement SOAs are promising because of their compactness and simple current pumping. The gain in a SOA saturates as the optical power level increases [6]. Singh *et al.* [7] presented the placement of semiconductor optical amplifier for 10Gbps non-return to zero format in dispersion–compensated fiber and single mode fiber link. In this work, different pre, post and symmetrical compensation methods for different locations of the SOA in fiber link have been reported. The impact of increase in signal input power for these three power compensation methods are compared in terms of BER, eye closure penalty and output received power. It was observed that the post power compensation method is better to symmetrical and pre power compensation methods when SOA is used. Woosuk Choi [8] analyzed the performance of 8 × 10 Gb/s transmission over 240 km wavelength division multiplexing system caused by crosstalk in cascaded conventional semiconductor optical amplifiers. Sun *et al.* [9] investigated an error-free transmission 32 × 2.5 Gb/s DWDM channels over 125 km using cascaded in-line SOAs. Kim et al. [10] successfully transmitted 10Gb/s optical signals over 80km through SSMF using SOAs as booster amplifiers. They have further found the suitable parameters of input signals for SOAs, such as rising/falling time, extinction ratio and chirp parameter to maximize output dynamic range and maximum output power.

In this paper, we extended the previous work by increasing the number of channels and transmission distance without any power compensation methods by using DS-Normal fiber only and investigated the performance of SOA for dense wavelength division multiplexed system by varying the confinement factor.

This paper is structured as follows. Section I presents introduction. Section II discusses the schematic setup of DWDM system using SOA. In Section III, results of SOA by varying confinement factor are presented and Section IV gives a brief outlook for the conclusion.

II. System Setup

Figure 1 is a system setup of an optical transmission system using semiconductor optical amplifier. The system consists of 32 channels whose wavelengths ranged from 192.35 THz-193.85 THz with a spacing of 100 GHz. Each transmitter consists of data source, laser source, NRZ rectangular driver, and optical amplitude modulator. Data source generates a pseudorandom binary sequence of data stream. The data is converted into optical signal using continuous wave lorentzian (CW) laser source Data format of the type NRZ rectangular is generated by the modulation driver. The pulses are modulated using Sin2 Mach-Zehnder modulator at 10 Gbps. The combined signals with an optical combiner were modulated simultaneously at 10 Gb/s. The signal was then input to DS-Normal fiber of the reference frequency 192.35 THz, dispersion-2ps/nm/km and attenuation is 0.2 dB/km. Then the output is fed to semiconductor optical amplifier through an optical splitter. The optimized parameters of SOA used in the simulation are as follows: bias current is 290mA, the length is 400 mm, the width of the active layer is 2 mm, and its thickness is 0.2 mm. The transparency carrier density in the SOA is taken to be 1.5×1018cm^3, Spontaneous Carrier lifetime is 0.3 ns and the differential gain 3 × 10-16 cm^2. The input and output coupling losses of SOAs are taken as 3 dB. The optical splitter (S1) is used to measure the optical input power for the transmission link. Optical Power meter and Optical probe with splitter (S2) are used for the measurement of optical signal power and spectrum at different levels.

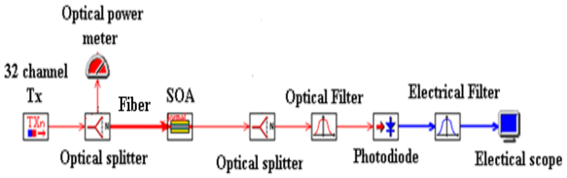


Fig. 1 System Setup based on Semiconductor Optical Amplifier

A single receiver consists of PIN photodiode, optical and electrical raised cosine filter. Optical filter component composed of a raised cosine transfer function filter having 1 as raised cosine exponent, 0.2 raised cosine roll off, 193.15 THz center frequency and band pass filter synthesis. the optical signal is converted into electrical signal using PIN photodiode. The PIN detector having responsivity of 0.87 and quantum efficiency 0.7 A/W. Electrical filter at the receiver side is raised cosine filter and has bandwidth of 8GHz. Electrical scope is used to attain the eye diagram. From the eye diagram, the values of Quality factor, BER and Eye closure can be analyzed.

III. Results & Discussion

The semiconductor optical amplifier has been investigated for 320 Gbps DWDM system in the term of quality of signal, bit error rate, eye closure and output power at channel-1. The Q factor vs. transmission distance at different confinement factors is as shown in Fig. 2. For 0.4 confinement factor, maximum transmission distance 245 km is covered and at 0.1 confinement factor, acceptable Q factor is obtained upto 105 km distance at -40 dBm signal input power. It is evident that the quality of the signal decreases with increasing the length of the fiber at low confinement factor. Our results are in coincidence with our previous results [5, 7] where we analyzed the performance of WDM system based on SOA with the same bit rate and channel spacing. Fig.3 indicates the bit error rate vs. transmission distance at different confinement factors. At 0.1 confinement factor, acceptable bit error rate (10–9) is achieved up to 105km. For 0.2 & 0.3 confinement factor, transmission distance increases up to 157 km & 201 km respectively. The transmission upto 245 km is achieved with acceptable bit error rate at 0.4 confinement factor. Fig.4 depicts the eye closure vs. transmission distance for different confinement factors at-40 dBm signal input power. We observed that 0.4 confinement factor provides least eye closure 2.13 dBm at 245 km distance. Means as we increase the transmission distance, the eye closure goes on increasing. As the eye closure increases, the quality goes on decreasing.

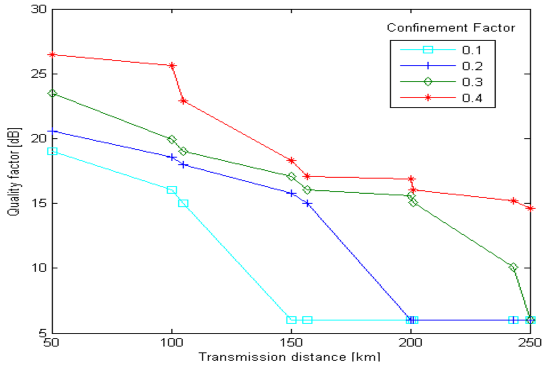


Fig. 2 Quality Factor vs. Transmission Distance for different Confinement Factors at -40 dBm Signal Input Power

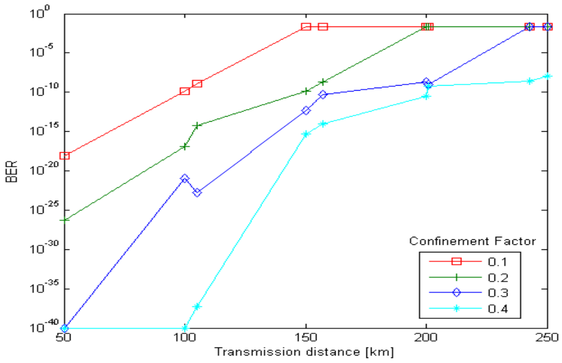


Fig. 3 BER vs. Transmission Distance for different confinement Factors at -40 dBm Signal Input Power

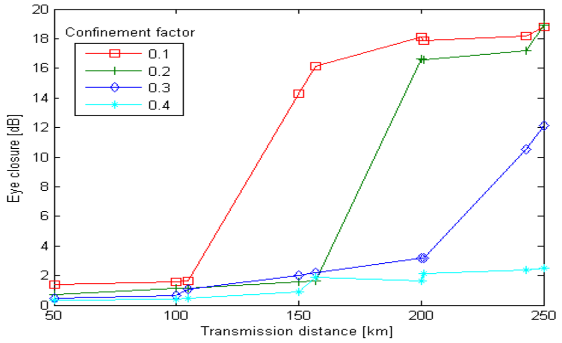


Fig. 4 Eye Closure vs. Transmission Distance for different Confinement Factors at -40 dBm Signal Input Power

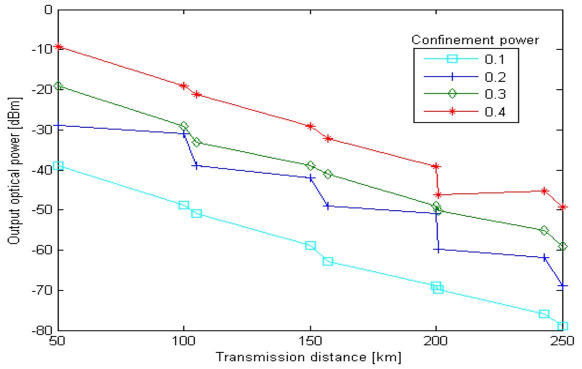


Fig. 5 Output Optical Power vs. Transmission Distance for different Confinement Factors at -40 dBm Signal Input Power

Figure 5 depicts the received signal power vs. transmission distance for different confinement factors. For 0.4 confinement factor upto 245 km transmission distance -46.16 dBm output power with the signal input power of -40 dBm is obtained.

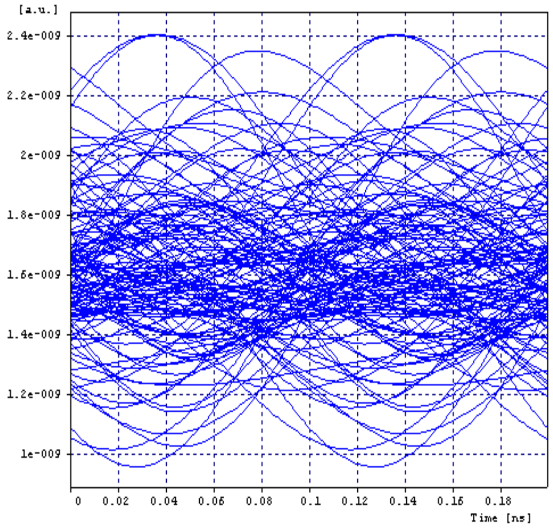


Fig. 6 Eye Diagram at 0.1 Confinement Factor for 215km Transmission Distance

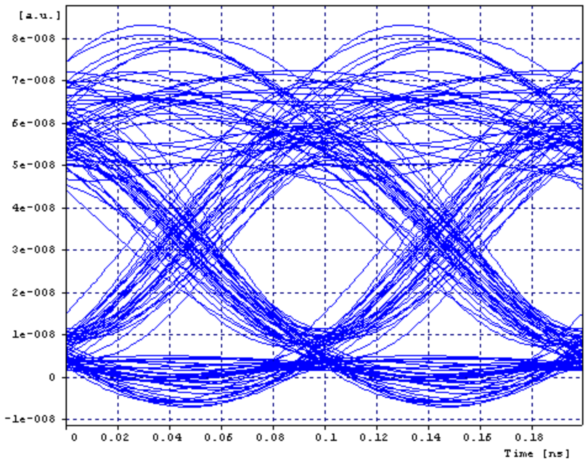


Fig. 7 Eye Diagram at 0.4 Confinement Factor for 215km Transmission Distance

The eye diagrams for 0.1 & 0.4 confinement factors are shown in Fig. 6 & 7. At low confinement factor of 0.1, acceptable BER and quality factor is obtained upto 105 km transmission distance. At low value of confinement factor the output signal is degraded. By increasing the confinement factor, large transmission distance is achieved. The 215km transmission distance is covered successfully for 0.4 confinement factor.

IV. Conclusion

In this paper, we have analyzed the performance of the 32 channel DWDM system based on optimized SOA at 10 Gbps with 0.8 nm channel spacing. The results have been reported for different confinement factors. We have observed that the signal can be transmitted successfully with improved performance upto 245 km transmission distance for 0.4 confinement factor of SOA at low input signal power of -40 dBm. For low confinement factor of 0.1, only 105km transmission distance is covered. In addition, it is found that the system provides acceptable performance up to 157 km & 201 km for confinement factors of 0.2 and 0.3. It is evident that the quality of the signal decreases at low confinement factors.

Acknowledgment

The authors would like to thank the Rsoft design group for the OptSim simulation software for optical communications systems.

References

[1] Y. Sun, A. K. Srivastava, S. Banerjee, J. W. Sulhoff, R. Pan, K. Kantor, R. M. Jopson, and A. R. Chraplyvy, “Error-free transmission of 32×2.5 Gbit/s DWDM channels over 125 km using cascaded in-line semiconductor optical amplifiers,” Electron. Lett., vol. 35, pp. 1863–1865, 1999.

[2] L. H. Spiekman, J. M. Wiesenfeld, A. H. Gnauck, and L. D. Garret, “Transmission of 8 DWDM channels at 20 Gb/s over 160 km of standard fiber using a cascade of semiconductor optical amplifiers,” IEEE Photon. Technol. Lett., vol. 12, pp. 717–719, June 2000.

[3] H. K. Kim, S. Chandrasekhar, A. Srivastava, C. A. Burrus, and L. Buhl, “10 Gbit/s based WDM signal transmission over 500 km of NZDSF using semiconductor optical amplifier as the in-line amplifier,” Electron. Lett., vol. 37, pp. 185–187, 2001.

[4] Christian Bohémond, Thierry Ramp one, and Ammar Sharaiha, “Performances of Photonic Microwave Mixer Based on Cross-Gain Modulation in a Semiconductor Optical Amplifier” Journal of Light wave Technology, vol. 29, no.16, pp. 2402-2409, August 15, 2011.

[5] Aruna Rani, Mr. Sanjeev Dewra, “Performance evaluation of DWDM system using Semiconductor optical amplifier in the presence of fiber non-linearities”, International Journal of Enhanced Research in Science Technology & Engineering, vol. 3 Issue 2, pp.301-305, February-2014.

[6] Bobby Barua, “Evaluate the Performance of Optical Time Division Demultiplexing with the Gain Saturation effect of Semiconductor Optical Amplifier” IACSIT International Journal of Engineering and Technology, Vol. 3, No. 5, October 2011

[7] Surinder Singh, R.S. Kaler “Placement of optimized semiconductor optical amplifier in Fiber optical communication systems” Optik, vol.119, pp.296–302, 2008.

[8] W. Choi, S. Hur, J. Lee, Y. Kim, J. Jeong, Transmission performance analysis of 8×10 Gb/s WDM signals using cascaded SOAs due to signal wavelength displacement, J. Lightwave Technol., vol. 20, pp. 1350–1356, 2002.

[9] Y. Sun, A.K. Srivastava, S. Banerjee, J.W. Sulhoff, Error free transmission of 32×2.5 Gbit/s DWDM channels over 125km using cascaded inline semiconductor optical amplifiers, Electron. Lett., vol. 35, pp.1863–1865, 1999.

[10] Y. Kim, H. Jang, Y. Kim, J. Lee, D. Jang, J. Jeong “Transmission performance of 10Gb/s 1550-nm transmitters using semiconductor optical amplifiers as booster amplifiers” IEE Journal of Lightwave Technology, Vol. 21, No.2, pp.476–481, 2003.

Analysis of 64 × 10 Gbps Dense Wavelength Division Multiplexing System Using Optimized RAMAN-EDFA Hybrid Optical Amplifier

Garima Arora[1] and Sanjeev Dewra[2]
[1,2]*Department of Electronics & Comm. Engineering,*
Shaheed Bhagat Singh State Technical Campus, Ferozepur–152004, Punjab, India
E-mail: [1]garima506@gmail.com, [2]Sanjeev_dewra@yahoo.com

***Abstract*—In this paper, the performance of 64 × 10 Gbps dense wavelength division multiplexing system using RAMAN-EDFA hybrid optical amplifier with 0.8 nm channel spacing has been demonstrated and the effect of Input signal power and bit rate has been investigated. The results have been carried out by evaluating the value of quality factor, BER and average opening of eye. It is found that using RAMAN-EDFA hybrid optical amplifier; the signal can be transmitted up to transmission distance of 135 km at low signal input power of-20 dBm with acceptable BER and Q-factor respectively. In addition, the system provides acceptable performance up to 165 Km and 200 Km with the signal input power of-10 dBm and 0 dBm.**

Keywords: Bit Error Rate, DWDM, RAMAN-EDFA

I. Introduction

The demand for broadcast over the global telecommunication network will continue to rise at an exponential rate and fiber optical system will be able to meet up the challenge. Optical fibers are considered as the main wire line physical medium for transmitting multimedia communication applications that require large amount of bandwidth [1]. Current efforts have been aimed at realization of greater capacity utilization of optical networks by multiplexing huge number of channels. These systems are referred to as dense wavelength-division multiplexing system. Dense WDM technology is recognized for its increase in transmission capability in optical communication system [2]. The combination of a multiple wavelengths on a single fiber allows for sharing of network elements such as amplifiers resulting in considerable cost savings [3]. The advancement in the fiber optic networks has been promoted by development of efficient optical amplifiers which eradicate the need of costly Optical-Electrical-optical conversions [4]. To extend the bandwidth and reach in optical communication system, the amplifiers having broadband and low-noise properties at reasonable cost are required. The high power conversion efficiency of erbium-doped fiber amplifiers [5] and broadband tunability combined with low-noise properties of Raman amplifiers [6] are employed in hybrid amplifier configurations to achieve better performance in WDM networks. The Hybrid amplifiers have wide gain bandwidth, more flat gain profile and high power gain which makes it advantageous over the conventional amplifiers [7].

Ju Han Lee *et al.* [8] investigated a new scheme of the dispersion-compensating Raman/EDFA hybrid amplifier recycling residual Raman pump for enhancement of overall power conversion efficiency. The significant improvement of gain of signal and gain bandwidth by 15 dB and 20 nm, compared to the performance of the Raman only amplifier was achieved by using this proposed scheme.

Singh *et al.* [9] evaluated the symmetrical, pre and post power compensation schemes for a dissimilar position of hybrid optical amplifier (RAMAN–EDFA) in an optical link. It was observed that the post power compensation scheme is better than symmetrical and pre power compensation schemes. Also the post power compensation scheme with −15 dBm signal input provides least BER of 10−40 and output power of 12 dBm.

R.S. Kaler [10] evaluated that the hybrid optical amplifier provides better performance when the optimized parameters are used. Hybrid optical amplifier (Raman-EDFA) had been optimized using different optimized parameter such as O/P power for fixed output power EDFA, noise figure and Raman pump, pump power, Raman fiber length for Raman amplifier. Further he found the maximum single span length using optimized hybrid optical amplifier. It was shown that using optimized hybrid optical amplifier, the dispersions at 2, 4, 8, and 16 ps/nm/km achieves 150, 150, 120 and 70 km of single span length with the acceptable BER.

Jian-guo Yuan *et al.* [11] introduced the configuration of the hybrid amplifier (Raman Amplifier: RA + Erbium-Doped Fiber Amplifiers: EDFA) and analyzed the restriction conditions of its optimum design. The ASE noise and Rayleigh noise in Raman Amplifier (RA) as well as their influences on the Signal Noise Ratio (SNR) of the receiver had been analyzed in depth. Furthermore, the influences which result from these noises on the performances of the optimum design for the hybrid amplifier (RA + EDFA) had thoroughly been researched.

The results presented in [9] for RAMAN-EDFA hybrid optical amplifier is carried out for a single channel. In this paper, we extended the previous work by increasing the number of channels with an improved power level and increase in maximum single span distance.

The paper is prepared in four sections. In Section 2, the system model is described. The simulation results have been calculated for the different input signal power and bit rates in section 3 and finally in Section 4, the conclusions are made.

II. Simulation Setup

The system model of 64 channel DWDM transmission system using RAMAN-EDFA hybrid optical amplifier with NRZ encoding technique and 0.8 nm interval is shown in fig. 1. An optical transmission link consists of three sections i.e. transmitter, optical amplifier and receiver. At the transmitting side, 64 channels are transmitted operating at its own frequency in range from 189.15 THz to 195.45 THz. Every transmitter consists of electrical driver, data source, laser source and external MZ modulator. The signal of 10 Gbps with pseudo random sequence is generated by the data source. Electrical driver generates Non Return to Zero rectangular type data format with a signal dynamics i.e. low level −2.5 and high level +2.5. The external MZ modulator receives the signals from data source and laser. The combiner is used to combine the all modulated optical signals and these signals are boosted by booster and fed into the optical fiber through an optical splitter. The optical splitters are used for the measurement of the optical power and to analyze optical spectrum the transmission link. The optical signals pre-amplified by booster are transmitted over DS anomalous fiber having reference frequency of 193.414 THz and attenuation of 0.2 dB/km.

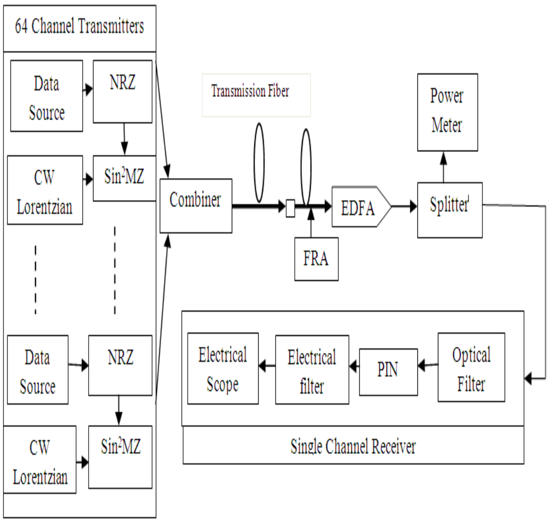


Fig. 1 System Model

Different components have different operational parameters. The type of EDFA used for amplification is fixed output power EDFA and its parameters are flat gain shape, fixed output power of 35mW and noise figure of 4.5 dB. The optimized parameters of Raman fiber are shown in Table 1.

Table 1 Optimized Parameters of Raman Fiber

Parameter	Value
Raman fiber length	10 Km
Operating temperature	300 K
Pump power	500mW
Pump Wavelength	1453 nm
Pump attenuation	0.2dB/km

The Optical raised cosine filter, PIN photodiode and Electrical raised cosine filter constitutes a single receiver. Raised Cosine Optical filter has band pass filter synthesis, 0.2 raised cosine roll off, 1 as raised cosine exponent and 40 GHz BW. The optical signal is converted to electrical signal using PIN photodiode, the Quantum efficiency of which is 0.7 A/W and responsivity is 0.87. Electrical raised cosine filter at the receiver side has bandwidth 8 GHz. The eye diagram is obtained from the Electrical Scope. The values of Quality factor, average opening of eye and Bit Error Rate can be analyzed from the eye diagram.

III. Results and Discussions

To evaluate the system performance, the eye diagram is analyzed for the first channel (189.15 THz) received at the receiver. Performance of 640 Gbps (64×10) dense Wavelength multiplexed system using RAMAN-EDFA hybrid optical amplifier is investigated by varying the distance from 50 to 250 Km. For different signal input power, the Q factor vs. transmission distance graph is as shown in Fig. 2. The quality of the received signal decreases as we increase the length of the fiber from 50 to 250 Km. The Quality factor varies as 21.14 to 6.02 dB for-20 dBm, 22.81 to 6.02 dB for -10 dBm and 24.74 to 6.02 dB for 0 signal input power. The acceptable Q factor is obtained up to 135 km transmission length for -20 dBm signal input power. For -10 dBm & 0d Bm signal input power acceptable quality factor is achieved up to 165km and 200km transmission length without dispersion compensation methods. This shows better results over [9] where the transmission length of 126 km was achieved with one channel using post power compensation scheme.

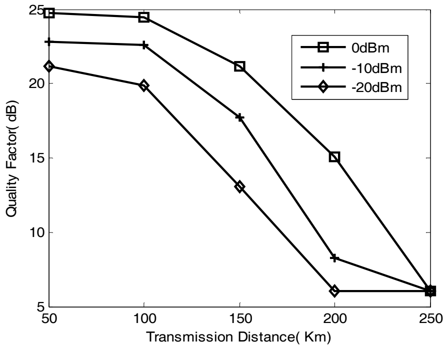


Fig. 2 Quality Factor vs. Transmission Distance for different Signal Input Power

Figure 3 reveals the plot of Bit Error Rate vs. transmission distance for different signal input power. The variation in BER from different input signal powers are 4.25 e -30 to 0.0227 for -20 dBm, 1.19 e -40 to 0.0227 for -10 dBm and 1 e -40 to 0.0227 for 0 input signal power. It is found that with an increase in the signal input power at the optical fiber link, the bit error Rate decreases.

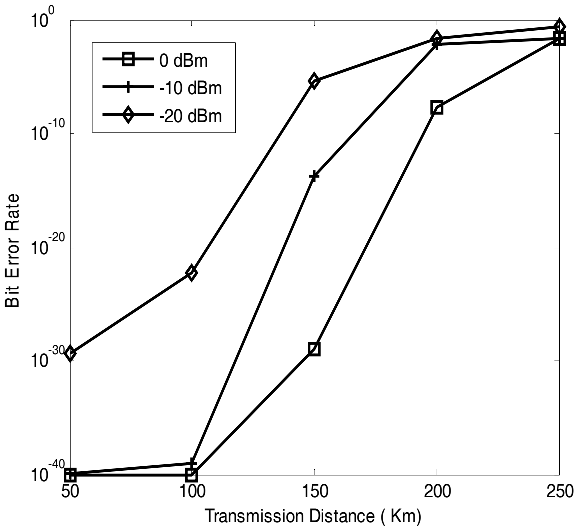


Fig. 3 Bit Error Rate vs. Transmission Distance for different Input Signal Power

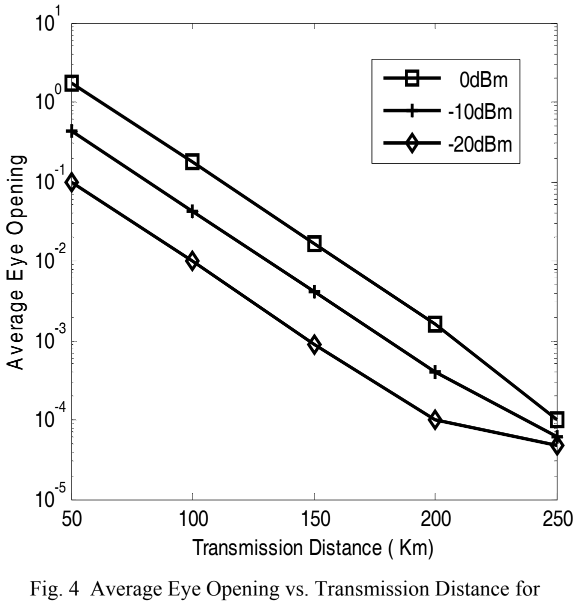


Fig. 4 Average Eye Opening vs. Transmission Distance for different Signal Input Power

The average eye opening from different input signal powers vs. transmission distance is shown in Fig. 4. It is evident that large opening of eye means less Bit Error Rate and high quality communication. It is found that average eye opening is decreasing, as the transmission length is increasing from 50 to 250 km. The average eye opening varies as is 0.0981 to 4.82 e -05 for -20 dBm, 0.4291 to 6.14 e -005 for -10 dBm and 1.7148 to 0.0001 for 0 dBm input signal power.

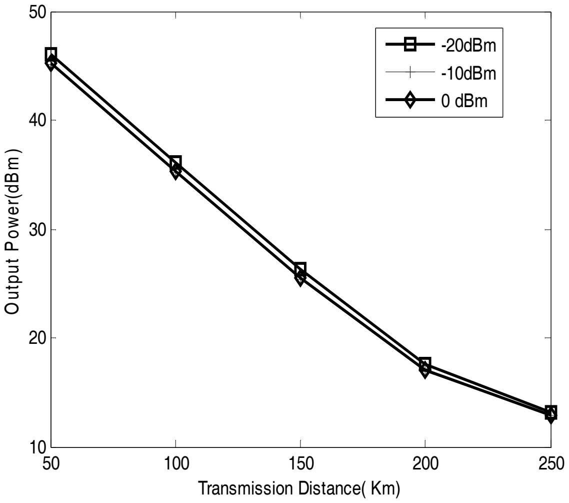


Fig. 5 Output Power vs. Transmission Distance for different signal input power

Figure 5 shows the output power vs. transmission distance. For-20dBm signal input power, the system shows the output power of 33.423 dBm at acceptable transmission distance of 135 Km which is an improvement over [9] in which the output power of 12dBm was obtained with the input signal power of-15 dBm at 126 Km with only one channel at input.

Further the performance of DWDM system was evaluated by varying the bit rates from 5 to 15 Gbps & has been evaluated on the basis of value of Q factor, Bit Error Rate and Average eye opening. The Quality Factor vs. transmission distance graph for different bit rates is shown in Fig. 6

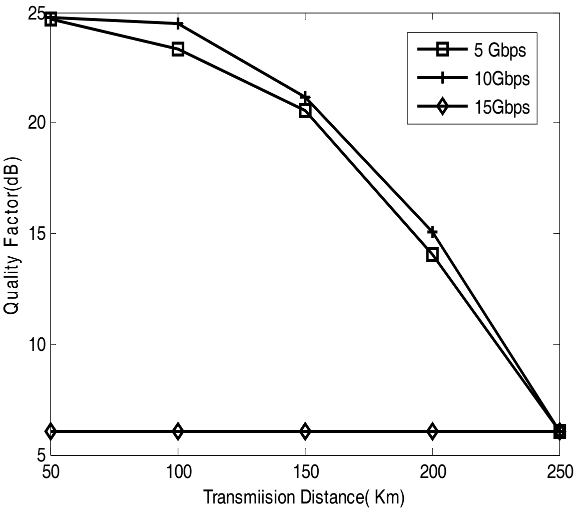


Fig. 6 Quality Factor vs. Transmission Distance at different Bit Rates

Figures 7 and 8 show the Bit Error rate and average eye opening vs. transmission distance for three different bit rates. The simulation results show that worst results are obtained for the bit rate of 15 Gbps.

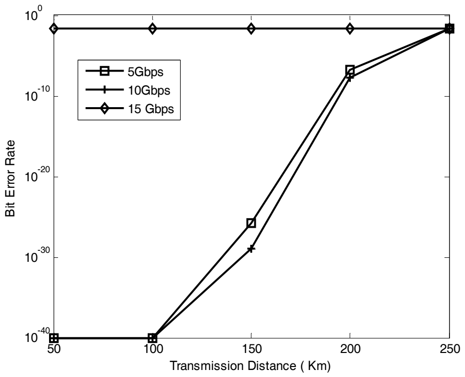


Fig. 7 Bit Error Rate vs. Transmission Distance at different Bit Rates

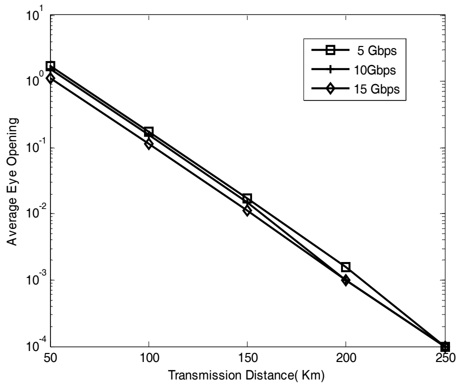


Fig. 8 Average Eye Opening vs. Transmission Distance at different Bit Rates

The eye diagram of signal after RAMAN-EDFA hybrid optical amplifier at 135 km distance with signal input power of-20dBm and bit Rate of 10Gbps is shown in Fig. 9

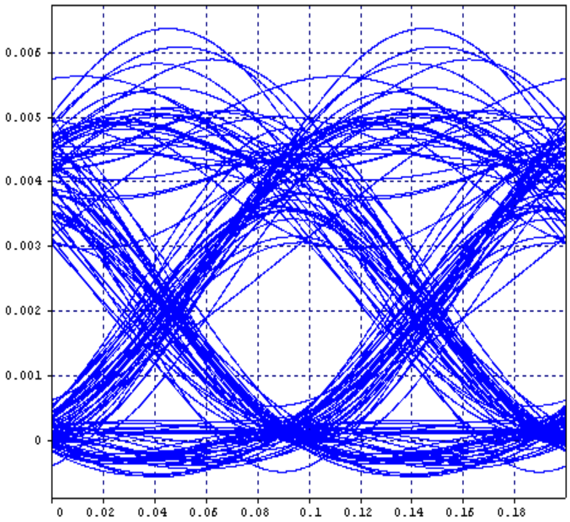


Fig. 9 Eye Diagram for RAMAN-EDFA at 135 Km with Signal Input Power of-20 dBm and Bit Rate of 10Gbps

IV. Conclusion

The 64 channel Dense Wavelength Division Multiplexing system at 10Gbps and 0.8 nm interval has been investigated with the RAMAN-EDFA hybrid optical amplifier. The performance was calculated for different bit rates and input signal power & has been evaluated on the basis of value of Q Factor, output power level, average eye opening and BER. The worst results are obtained at the bit rate of 15Gbps. Further, it is observed that the RAMAN-EDFA hybrid optical amplifier provides acceptable quality factor (15.57dB) and bit error rate (1.04e-09) up to 135 km transmission distance for-20dBm signal input power with without any dispersion compensation methods.

References

[1] Mustafa M. Matalgah and Redha M. Radaydeh, “Hybrid Frequency-Polarization Shift-Keying Modulation for Optical Transmission” IEEE journal of lightwave technology, Vol. 23, no. 3, March 2005.

[2] C.A. Brackett, “Dense wavelength division multiplexing networks: Principles and applications,” IEEE J. Sel. Areas Commun. 8, pp. 948–964, 1990.

[3] Sanjeev Dewra, R. S. Kaler, “ Performance evaluation of an optical network based on optical Cross add drop multiplexer” J. Opt. Technol. 80 (8), August 2013.

[4] Garima Arora, Mr. Sanjeev Dewra, “DWDM Transmission using Hybrid Optical Amplifiers” International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 4, April 2014..

[5] A. Ahmad et al., Investigation of hybrid gain-clamped Raman-fiber amplifier/ EDFA utilizing pump reuse technique, Laser Phys. Lett. 5, pp no. 202–205, 2008.

[6] C. Sun Hyok et al., Characteristics of low noise hybrid fiber amplifier, Opt. Commun. 261, pp no. 269–275, 2006.

[7] Garima Arora, Mr. Sanjeev Dewra, “A review on optical Amplifier its Hybrid Configurations, International Journal of Electronics & Communication Engineering Research” Vol. 1 Issue 6, November-2013.

[8] Ju han Lee, “dispersion compensating RAMAN/EDFA hybrid amplifier recycling residual RAMAN pump for efficiency enhancement” IEEE photonics letters, Vol 17, no. 1, 2005.

[9] Simranjit Singh, R.S. Kaler, “Placement of hybrid optical amplifier in fiber optical Communication systems” Optik 123, pp. 1636–1639, 2012.

[10] R.S. Kaler, “Optimization of hybrid Raman/erbium-doped fiber amplifier for multi terabits WDM system” Optik 124, pp no. 575– 578, 2013.

[11] Jian-guo Yuan, Tian-yu Liang, Wang, Sheng Gu, “Impact analysis on performance optimization of the hybrid amplifier (RA + EDFA)” Optik 122 pp no. 1565– 1568, 2011.

Performance Analysis of VMM Using LINPACK and CLOUDSIM

Bohar Singh, Pawan Luthra and Bindu Bala
[1,2]*Computer Science and Engineering,*
Shaheed Bhagat Singh State Technical, Campus, Ferozepur, India
[3]*Computer Application Department,*
Shaheed Bhagat Singh State Technical, Campus, Ferozepur, India
E-mail: [1]*Bohar2@gmail.com,* [2]*Pawanluthra81@gmail.com,* [3]*Dotmca@gmail.com*

***Abstract*—Virtualization is a core part of cloud computing which divides the resources of a computer into multiple execution environments. Virtualization offers a lot of benefits including flexibility, security, ease to configuration and management and reduction of cost but at the same time it also brings a certain degree of performance overhead. Furthermore, Virtual Machine Monitor (VMM) is the core component of virtual machine (VM) system and its effectiveness greatly impacts the performance of whole system. In this paper, different software & tools i.e VMware, VirtualBox, LINPACK Benchmark and CloudSim are used and conducted experiments to measure the performance of virtualized XP are described. The quantitative and qualitative comparison of both virtual machine monitors can be done by measuring the processing speed and response time of virtualized XP.**

Keywords: LINPACK, CloudSim, Virtual Machine Monitor (VMM), VMware and VirtualBox

I. INTRODUCTION

With the rapid development of internet, Cloud computing has recently emerged as a new paradigm for hosting and delivering services over the Internet [1,2]. Cloud computing is best for business owners to grow from lower resources to large due to increasing demand of service. This technological trend has enabled the realization of a new computing model called cloud computing, in which user can access services from anywhere, at anywhere over the Internet in an on-demand fashion [2].

The deployment of clouds have many attractive benefits such as scalability and reliability, reduced cost, elasticity, quality of service (QoS); however, development of clouds to provide more economical solutions as consumers only need to pay for what resources they required while providers can capitalize poorly utilized resources. From a provider's point of view, the maximization of the profit is major concern.

Many years ago, a problem invoked that how to run multiple operating systems on the same machine simultaneously? The solution to this problem was virtual machines created by VMM [1,8]. Virtual machines monitor the core part of virtual machines sits between one or more operating systems and the hardware and gives the illusion to each running OS that it controls the machine. Virtual Machine technology starts to focus on virtualization [9,10] which enables to run multiple OS environments simultaneously on the same physical hardware in strong isolation. Virtualization can be used to perform isolation, consolidation and migration of resources in cloud computing. VMs share the conventional hardware in a secure manner with help of VMM which control guest OS on host, while each VM is hosting its own operating system and applications.

Virtual Machine technology [8, 9] begins to emerge as a focus of research and deployment. Virtual Machine technology such as Xen, VMWare, Microsoft Virtual Servers, and new Microsoft Hyper-V technology etc, enables multiple OS environments to coexist on the same physical computer in strong isolation. VMs share the conventional hardware in a secure manner with excellent resource management capacity, while each VM is hosting its own operating system and applications. Hence, VM platform can facilitate server-consolidation [11] and co-located hosting facilities.

Behind the scenes, the monitor actually is in control of the hardware, and must multiplex running OSes across the physical resources of the machine. Indeed, the VMM serves as an operating system for operating systems but at a much lower level; the OS must still think it is interacting with the physical hardware. Thus, transparency is a major goal of VMMs [5].

At the end of the 1960s, the virtual machine monitor (VMM) came into being as a software abstraction layer that partitions a hardware platform into one or more virtual machines. Each of these virtual machines was sufficiently similar to the underlying physical machine to run existing software unmodified. There was time; general-purpose computing was the domain of large, expensive mainframe hardware and users found that VMMs provided a compelling way to multiplex such a scarce resource among multiple applications. Thus it brought modern multitasking operating systems and a simultaneous drop in hardware cost, which eroded the value of VMMs. As mainframes gave way to minicomputers and PCs, VMMs disappeared because computer architectures no longer provided the necessary hardware to implement them efficiently.

By the late 1980s, neither academics nor industry practitioners viewed VMMs as much more than a historical curiosity. Fast forwarding to 2005, VMMs becomes hot topic In academia and industry. Virtual-machine-based technologies such as Intel, AMD, Sun

Microsystems, and IBM are developing virtualization strategies that target markets with revenues in the billions and growing. In research labs and universities, researchers are developing approaches based on virtual machines to solve mobility, security, manageability and virtualization problems.

In the 1990s, Stanford University researchers began to look at the potential of virtual machines to overcome difficulties that hardware and operating system limitations imposed: This time the problems stemmed from massively parallel processing (MPP) machines that were difficult to program and could not runs existing operating systems. With virtual machines, researchers found they could make these unwieldy architectures look sufficiently similar to existing platforms to leverage the current operating systems. Today virtual machine has become popular due to various reasons. Server consolidation is one such reason. In many settings, people run services on different machines which run different operating systems and yet each machine is lightly utilized. In this case, virtualization enables an administrator to consolidate multiple Operating systems onto fewer hardware platforms and thus lower costs and ease of administration. Virtualization has also become popular on desktops, as many users wish to run one operating system but still have access to native applications on a different platform. This type of improvement in functionality is also a good reason. Another reason is testing and debugging. While developers write code on one main platform, they often want to debug and test it on the many different platforms that they deploy the software to in the field. Thus, virtualization makes it easy to do by enabling a developer to run many different operating system types and versions on just one machine [18,19].

II. VIRTUAL MACHINE MONITORS (VMM)

A. *VMware*

VMware [10] is a virtualization and cloud computing software provider for x86 compatible computer architecture. VMware Inc. is a subsidiary of EMC Corporation, well known in the field of system virtualization and cloud computing. VMware's software allows users to create multiple virtual environments, or virtual computer systems, on a single computer or server. Essentially, one computer or server could be used to host, or manage, many virtual computer systems, sometimes as many as one hundred or more. The software virtualizes hardware components such as the video card, network adapters, CPU, memory and hard drive called hardware virtualization which is very useful for enterprisers by setting up multiple server systems on single server or physical hardware without having to purchase separate hardware for each of them. User can create virtual servers using VMware's software, leads to saving a lot of time and money [11].

Virtual machine (VM) is heart of virtualization, tightly isolated software container with an operating system and application inside. Because each virtual machine is completely separate and independent, many of them can run simultaneously on a single computer. thin layer of software called a hypervisor decouples the virtual machines from the host and dynamically allocates computing resources to each virtual machine as needed.

The x86 architecture offers four levels of privilege known as Ring 0, 1, 2 and 3 to operating systems and applications to manage access to the computer hardware. While user level applications typically run in Ring 3, the operating system needs to have direct access to the memory and hardware and must execute its privileged instructions in Ring 0. Virtualizing of x86 architecture require placing a virtualization layer under the operating system to create and manage the virtual machines that deliver shared resources. Some sensitive instructions cannot effectively be virtualized as they have different semantics when they are not executed in Ring 0. The difficulty in trapping and translating these sensitive and privileged instruction requests at runtime was the challenge that originally made x86 architecture virtualization look impossible so VMware resolved this problem by developing binary translation techniques that allow the VMM to run in Ring 0 for isolation and performance, while moving the operating system to a user level ring with greater privilege than applications in Ring 3 but less privilege than the virtual machine monitor in Ring 0. It does not support Hyper Threading3 and requires a host operating system, which means an extra layer and additional overhead.

B. *VirtualBox*

VirtualBox is a cross-platform virtualization application, installs on the existing Intel or AMD-based computers, whether they are running Windows, Mac, Linux or Solaris operating systems. Secondly, it extends the capabilities of a user's existing computer so that it can run multiple operating systems inside multiple virtual machines at the same time. A user can run Windows and Linux on Mac, run Windows Server 2008 on Linux server, run Linux on Windows PC. User can install and run many virtual machines, but the only practical limits are disk space and memory.

VirtualBox is simple and easy to use, y*et al*so very powerful to manage VM. It can run everywhere from small embedded systems or desktop class machines. Oracle VirtualBox is an x86 cross platform open source virtualization software package developed by Oracle Corporation as part of its family of virtualization products. Virtual Box is also called hosted hypervisor, host OS is required on which VMM is installed. To a very large degree, VirtualBox is functionally identical on all of the host platforms and same file, image formats are used. It provides the option to enable hardware virtualization on a per virtual machine basis when running on AMD-V and Intel-VT capable CPUs. On more recent CPU designs, VirtualBox is also able to make use of nesting paging tables to improve virtual machine performance.

III. Benchmark

A. *LINPACK*

The LINPACK [3,4] package was based on another package, called the Level 1 Basic Linear Algebra Subroutines (BLAS) [5]. Most of the floating-point work within the LINPACK algorithms is carried out by BLAS, which makes it possible to take advantage of special computer hardware without having to modify the underlying algorithm [4, 5]. In the LINPACK Benchmark, a matrix of size 100was originally used because of memory limitations with the computers. Such a matrix has 10 000 floating-point elements and could have been accommodated in most environments of that time. At the time it represented a large enough problem.

LINPACK (LINear system PACKage) contains a number of FORTRAN subroutines based on BLAS (Basic Linear Algebra Subprograms) library to work out different linear equations and linear least-squares problems written by Jack Dongarra, Jim Bunch, Cleve Moler and Pete Stewart. LINPACK [7] was designed for applying to supercomputers in the 1970s and early 1980s and now acts as one of the most authoritative benchmarks in high performance computers. The TOP 500 computers in the world are sorted by the LINPACK's result. To follow the development of computer architectures, LINPACK evolves into EISPACK and LAPACK. EISPACK mainly dedicates to numerical computation of the Eigen values and eigenvectors of matrices. LINPACK measures the actual peak value of float-point computing power indicated in giga of float-point operations per second (GFLOP) [3-5].

The original LINPACK Benchmark is an accident and was originally designed to assist users of the LINPACK package by providing information on the execution times required to solve a system of linear equations. The first 'LINPACK Benchmark' report appeared as an appendix which comprises data for one commonly used path in the LINPACK software package. Results were provided for a matrix problem of size 100, on a collection of widely used computers. This was done so users could estimate the time required solving their matrix problem by extrapolation. The LINPACK package is a collection of FORTRAN subroutines for solving various systems of linear equations. The software in LINPACK is based on a de-compositional approach to numerical linear algebra. The package has the capability of handling many different matrix and data types and provides a range of options.

B. *CloudSim*

CloudSim is a new generalized and extensible simulation toolkit and application which enables seamless modelling, simulation, and experimentation of emerging cloud computing system, infrastructures and application environments for single and internetworked clouds. The existing distributed system simulators were not applicable to the cloud computing environment due to evaluating the performance of cloud provisioning policies, services, application workload, models and resources under varying system, user configurations and requirements. To overcome this challenge, CloudSim can be used. In simple words, CloudSim is a development toolkit for simulation of Cloud scenarios. CloudSim is not a framework as it does not provide a ready to use environment for execution of a complete scenario with a specific input. Instead, users of CloudSim have to develop the Cloud scenario it wishes to evaluate, define the required output, and provide the input parameters. CloudSim is invented as CloudBus Project at the University Of Melbourne, Australia and supports system and behaviour modelling of cloud system components such as data centers, virtual machines (VMs) and resource provisioning policies. It implements generic application provisioning techniques that can be extended with ease and limited efforts. CloudSim helps the researchers to focus on specific system design issues without getting concerned about the low level details related to cloud-based infrastructures and services. CloudSim is an open source web application that launches preconfigured machines designed to run common open source robotic tools, robotics simulator Gazebo. SimJava is a toolkit for building working models of complex systems. It is based around a discrete event simulation kernel at the lowest level of CloudSim. It includes facilities for representing simulation objects as animated icons on screen.

IV. Problem Formulation

Cloud computing is emerging as a significant shift as today's organizations which are facing extreme data overload and high energy costs. Many Years ago, a problem aroused. How to run multiple operating systems on the same machine at the same time? The solution to this problem was virtual machines. Virtual machines monitor, the core part of virtual machines sits between one or more operating systems and the hardware, it also gives the illusion to each running OS that controls the machine. Behind the scenes, the monitor actually is in control of the hardware, and must multiplex running OSes across the physical resources of the machine. Indeed, the VMM serves as an operating system for operating systems, but at a much lower level; the OS must still think it is interacting with the physical hardware. Thus, transparency is a major goal of VMMs.

V. Experiment Setup

Today, the benchmark is used by scientists and engineers worldwide to evaluate computer performance, particularly for innovative advanced-architecture machines. An accurate performance evaluation is a complex issue so to accommodate its evaluation; the LINPACK Benchmark suite provides separate benchmarks that can be used to evaluate computer performance on a dense system of linear equations for a 100 × 100 matrix and is dependent on the algorithm chosen by the manufacturer and the amount of memory available on the computer being benchmarked. In the case of LINPACK 100, the problem size was relatively small and no changes were allowed to the LINPACK software.

Thus, as described before, many high performance machines may not have reached their asymptotic execution rates. However, the benchmark is still important because it approximates the performance rates of numerically intensive codes written by the user and optimized by an optimizing compiler quite well. The comparison of VMware and VirtualBox is done by benchmarking virtual machines using LINPACK and CloudSim. The virtual machines were created on VMware and VirtualBox One on VMware and another on VirtualBox each with the fixed configuration. Both contains Windows XP SP3 edition as guest OS. In the first experiment, LINPACK is used to compare the floating point operations of virtualized XP in VMware and virtualized XP. The input to the LINPACK tool was the problem size. The configurations were as follows:

Table 1 Experimental Configuration

Processor	Intel i3
RAM	1 GB
CPU Frequency	2.294 GHz
No. of CPU's	1
No. of Cache	2
No. of thread	4
No. of virtual machine	2

The output is the number of GFlop operations per second. 5 tests were performed with problem sizes 1000, 2000, 3000, 4000, 5000 their corresponding leading dimensions being 2000, 4000, 6000, 8000 and 10000 respectively. The number of trials will be 8 with data alignment value of 8 KB.

In the second experiment, CloudSim is used which is a cloud network simulator tool to compare the response of each virtual machine. The client sends the data to the server and CloudSim checks the speed and the quality of the networking. Here, we have the same system as the client and server. Server receives the packets from a port which is different from the port from where the client sent the data.

The simplest one consists of modelling the case were a single, centralized Cloud data center is used to host the social network application. In this model, all requests from all users around the world are processed by this single data center. Data center has 5 virtual machines allocated to different 5 applications. The second scenario consists of the use of two data centers, each one with 5 virtual machines dedicated to the application. The third scenario consists of three data centers, each one with 5 virtual machines. In the next scenario, the MIPS used by Cloudlet and VM are same or different. Each of these scenarios was evaluated with execution of the workload previously described. Results are discussed next.

VI. Results

The set of performance results reveals the asymptotic behaviour of VMM on a cluster of Intel i32.294 GHZ processors. Performance tests were also performed on a larger system. The LINPACK Benchmark results for this system are presented in Table 2: which shows LINPACK results on VMware and Table 3: shows LINPACK results on VirtualBox. However Fig 1 shows there is slightly more GFlops in case of VMware than VirtualBox with very small variation. So from this experiment it can conclude that VMware performs slightly better than VirtualBox.

Table 2 LINPACK's Results on VMware

Problem Size	Array Size	Trail	Time(ms)	GFlop
1000	2000	8	0.08	8.18
2000	4000	8	0.43	12.20
3000	6000	8	1.60	11.01
4000	8000	8	3.56	11.97
5000	10000	8	6.90	12.11

Table 3 LINPACK's Results on VirtualBox

Problem Size	Array Size	Trail	Time(ms)	GFlop
1000	2000	8	0.10	7.43
2000	4000	8	0.46	11.90
3000	6000	8	1.58	10.50
4000	8000	8	3.73	11.67
5000	10000	8	7.20	11.51

With respect to cloud simulation, the Table. 4 and 5 depicts variation of average response time. Results show that bringing the service closer to users improves the quality of service such as response time. It is an expected effect, because users experiment less effects from high latency and low bandwidth when they are geographically close to the application server. Results also show that service quality can be further improved with the application of load balancing in the application across data centers, which are supposed to be managed by different service brokerage policies and at virtual machine level. But the levels of improvement achieved depend largely on the load balancing algorithms employed. So, application of good load balancing strategies is paramount for large-scale distributed applications. It is not a matter of great concern in Cloud data centers, which apply economy of scale to make their business profitable and so they can offer more resources during peak traffic.

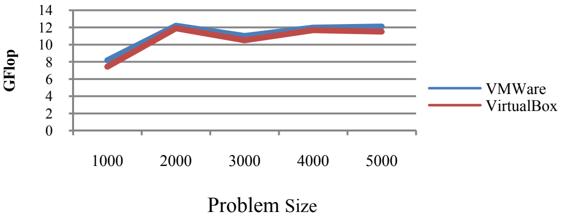


Fig. 1 Linpack Comparison Chart

TABLE 4 CLOUDSIM RESULTS ON VMWARE

Datacenter	VM	Host	Cloudlet	CL & VM	Response Time	Processing Time (ms)
1	1	1	1	Yes	0.2	1000
1	1	1	1	No	0.2	500
1	5	5	5	Yes	0.2	2000
1	5	5	5	No	0.2	500-2000
2	1	1	1	Yes	0.2	1000
2	1	1	1	No	0.2	2000
2	5	1	5	Yes	0.2	2000
2	5	1	5	No	0.2	500-1000
3	5	1	5	Yes	0.3	2000
3	5	1	5	No	0.3	500-1000

TABLE 5 CLOUDSIM RESULT ON VIRTUAL BOX

Datacenter	VM	Host	Cloudlet	CL & VM	Response Time	Processing Time (ms)
1	1	1	1	Yes	0.2	1000
1	1	1	1	No	0.2	500
1	5	5	5	Yes	0.2	2000
1	5	5	5	No	0.2	1500
2	1	1	1	Yes	0.2	1000
2	1	1	1	No	0.3	2000
2	5	1	5	Yes	0.3	2000
2	5	1	5	No	0.3	1000
3	5	1	5	Yes	0.3	2000
3	5	1	5	No	0.4	2000

Once again, elastic cloud providers solve this problem by charging consumers proportionally to the amount of resources used. At the same time provider offer tools to automatically increase and decrease resources available to applications in order to meet established service level agreements (SLA).

VII. CONCLUSION AND FUTURE SCOPE

The basic knowledge about VMware, VirtualBox, benchmark LINPACK, CloudSim and conducted experiments to measure the performance of virtualized XP on VMware and VirtualBox are described. The processing speed in the basis of number of floating point operations per second and the response time are measured. Significant results were produced with which the comparison can be done. The recent trends in high-performance computing will shape the near future of the LINPACK and CloudSim Benchmark suite.

REFERENCES

[1] J.H. Che, Q.M. He, Q. H. GAO and D. W. Huang, "Performance Measuring and Comparing of Virtual Machine Monitors", International Conference on Embedded and Ubiquitous Computing (EUC2008), 2008.

[2] Xianghua Xu, Feng Zhou, Jian Wan Yucheng Jiang, "Quantifying performance properties of virtual machine",International Symposium on Information Science and Engineering, 2008.

[3] J.J. Dongarra, P. Luszczek and A. Petitet. "The LINPACK Benchmark: past, present and future", Concurrency and Computation Practice and Experience, 15(9):803−820, 2003.

[4] Dongarra JJ, Bunch J, Moler C, Stewart GW, "LINPACK User's Guide", SIAM: Philadelphia, PA, 1979.

[5] Lawson C, Hanson R, Kincaid D, Krogh F., "Basic Linear Algebra Subprograms for Fortran usage", ACM Transactions onMathematical Software, Vol: 5, PP: 308–323, 1979.

[6] Dongarra JJ, "Performance of various computers using standard linear equations software", Technical Report CS-89-85, University of Tennessee, 2002.

[7] Edward Anderson, Z. Bai *et al.*," LAPACK User's Guide", Society for Industrial and Applied Mathematics, Philadelphia, PA, Third edition, 1999.

[8] M. Rosenblum and T. Garfinkel, "Virtual Machine Monitors: Current Technology and Future Trends", IEEE Computer, Vol: 38, Issue: 5, 2005.

[9] R. Uhlig, G. Neiger, D. Rodgers, *et.al.* "Intel Virtualization Technology", IEEE Computer, Vol: 38, Issue: 5, 2005.

[10] VMware Inc., "Understanding Full Virtualization, Paravirtualization and Hardware Assist", White paper, 2007.

[11] P. Padala, X. Zhu, Z. Wang, S. Singhal and K. Shin, "Performance Evaluation of Virtualization Technologies for Server Consolidation", Technical Report HPL-2007-59, HP Labs, April 2007.

[12] V. Inc., "A Performance Comparison of Hypervisors", Technical report, VMWare Inc., 2007.

[13] Jon dugan. Iperf Tutorial Summer JointTechs 2010, Columbus, OH

[14] P. Apparao, R. Iyer, X. Zhang, D. Newell and T. Adelmeyer, "Characterization & analysis of a server consolidation benchmark", Proceedings of 4th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments (VEE'08), USA, pages 21-30, 2008.

[15] R. Shiveley, "Enhanced Virtualization on Intel R Architecture-based Servers.Technology", 2005.

[16] Jinho Hwang, "A Component-Based Performance Comparison of Four Hypervisors",978-3-901882-50-0 copyright 2013, IFIP.

[17] L.H. Seawright and R. A. MacKinnon, "VM/370−A Study of Multiplicity and Usefulness", IBM Systems Journal. 18(1):4−17, 1979.

[18] Vasudevan.M.S, Biju.R.Mohan and Deepak.K.Damodaran, "Performance Measuring and Comparison of VirtualBox and VMware", International Conference on Information and Computer Networks (ICICN 2012), IPCSIT vol. 27 (2012)

[19] Bohar Singh, PawanLuthra"Review of Linpack and Cloudsim on VMM",International Journal of Engineering Trends and Technology (IJETT), (ISSN: 2231-5381) – Volume 11 Number 6-May 2014

[20] Jianhua Che, Qinming He, Qinghua Gao, Dawei Huang," Performance Measuring and Comparing of Virtual Machine Monitors", 2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, pages 381-386, 2008.

Application of Type-2 Fuzzy Logic – A Review

Satvir Singh[1], Inderjeet Singh Gill[2], Sarabjeet Singh[3] and Gaurav Dhawan[4]
[1,2]Department of Electronics & Comm. Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
[3,4]Department of Computer Sc. & Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: [1]drsatvir.in@gmail.com, [2]er.inderjeet@gmail.com, [3]sarabjeet_singh13@yahoo.com,
[4]dhawangaurav200@gmail.com

***Abstract*—In this paper, a comprehensive survey on various applications of Type-2 Fuzzy Logic has been carried out. GPU (Graphics Processing Units) is used for speed up purpose, otherwise commonly used for graphics applications. However, recent trends show use of GPU in various other general computational applications to run them parallel to reduce overall execution time. In this paper, it is discussed that how GPU and fuzzy logic can together be helpful in solving problems of different domains for faster responses.**

Keywords–GPU, CUDA, FLS, Type II Fuzzy Logic

I. INTRODUCTION

Latest advancements in parallel computing exploit GPGPUs that have multi-core architecture which supports parallel computations especially required for graphical processing. They devote more transistors for arithmetic and logical operations as compared to data caching and flow control compared to a CPU. Due to processing demand GPUs have advanced rapidly. And also beating the CPUs in terms of number of cores and hence, their computational power. As NVIDIA has launched CUDA software development kit in 2007, the use of GPU's computational powers for general purpose computing has become easy. It gives an API built upon the C language that can be used to write parallel computer programs. The GPU device operates as a coprocessor to the host i.e., CPU, running C program.

The paper is organized in six sections, in section II brief description of GPU is provided. In section III CUDA features are discussed, section IV and V provides information about fuzzy logic system and Type-2 fuzzy logic system and finally section VI provides various applications where Type-2 fuzzy logic can be used followed by conclusion.

II. GRAPHICS PROCESSING UNITS

GPU is a main hardware specially designed for highly parallel applications. The GPU's fast increase in both programmability and capability has spawned a research community that has successfully mapped a broad range of computationally demanding and complex problems to the GPU. This effort in general purpose computing with GPU, also known as GPU computing [1] [2]. GPUs have been known to users for quite a long time as a graphics rendering coprocessor to the host PC, to render cool graphic effects in multimedia based applications such as gaming, animation etc. But now the technology inside the GPU has became advanced for many computing applications other than rendering graphics. The research community has clearly demonstrated how non-graphics-related computing can be performed on the GPUs, with more than a thousands of papers published so far in this field. So GPGPU is use of GPU computing for general purpose applications. GPU is invented by NVIDIA.

GPU computing uses GPU together with a CPU to accelerate general-purpose scientific and engineering applications. CPU sends tasks and data to GPU, GPU performs computations on data and sends back results to CPU. GPU is called as DEVICE and CPU is called as HOST. GPUs consist of thousands of smaller, more efficient cores designed for parallel performance. CPUs consist of a few cores optimized for serial processing e.g. Intel Pentium Dual Core processors have 2 cores, Quad core have 4, which are very less in number. The architecture of CPU and GPU is shown in Fig. 1. Serial portions of the code run on the CPU while parallel portions run on the GPU. GPUs contain much larger number of dedicated ALUs then CPUs. Each processing unit on GPU contains local memory that improves data manipulation and reduces fetch time.

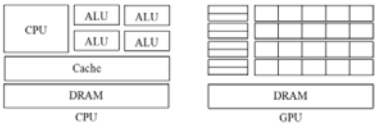


Fig. 1 Difference Between CPU and GPU Architecture [1]

III. CUDA

CUDA is NVIDIA's solution to access the GPU. Compute Unified Device Architecture (CUDA) is a data-parallel computing environment that does not require the use of a graphics API [2]. To work on CUDA, C language is being used. A CPU and a GPU programs are build up in the same environment i.e., CUDA-C language. In CUDA multiple kernels run concurrently on a single GPU. CUDA mention each kernel as grid. A grid consists of collection of blocks. Each block runs the same kernel but independent of each other. A block contains threads, which are smallest divisible unit on a GPU. CUDA allows multiple programs, kernels; to run sequentially on a single GPU [3]. The architecture is shown in Fig. 2.

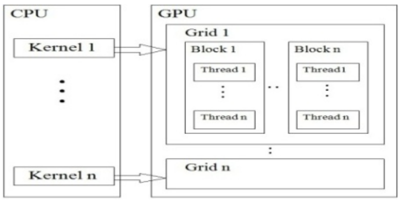


Fig. 2 CUDA Processing Model Design [1]

IV. FUZZY LOGIC SYSTEM

A Fuzzy Logic System (FLS) is able to handle the numerical data and linguistic knowledge simultaneously. FLS can be explained in form of mathematics a linear combination of fuzzy basis function and is a nonlinear universal function approximation. The fuzzy basis function expansion is very useful because its basis functions can be derived from either objective knowledge or subjective knowledge, both of which can be assigned into the forms of IF-THEN rules. Both type of knowledge can be expressed in mathematical manner. There are two types of problem knowledge which can be solved by the Fuzzy Logic System.

1. Objective Knowledge which is used for mathematical models. For example solve the formula, equations of motion for a submarine, spacecraft etc.
2. Subjective knowledge which represents linguistic information that is usually impossible to calculate value using mathematical formulas, the rules that might be valid for tracking a submarine or any other slowly moving large object etc.

FLS is a non-Linear mapping of an input data (feature) vector into a scalar output [4]. Fuzzy logic is a process that tries to simulate the “fuzzy” decision making of a person, by using fuzzy sets. Fuzzy logic often only requires a small number of fuzzy sets and a small collection of simple rules to solve the same problem. In fact, when dealing with a fuzzy problem, computers that operate using fuzzy logic often perform tasks more quickly, efficiently and in many cases better than normal computers which use traditional crisp logic. An example which illustrates the difference between fuzzy and crisp logic is the way in which a computer controls an air conditioner. The normal crisp-logic computer has a sensor that measures the temperature, after which this number is fed into a computer that has some built-in logical rules under which it operates. A fuzzy logic system has three main components:

1. Fuzzifier: A fuzzifier that takes in numbers (in this case the temperature) and transforms them into fuzzy sets.
2. Logic Control Center: A logic control center that uses rules, which are activated by fuzzy sets, and, produces fuzzy sets at its output.
3. Defuzzifier: A defuzzifier that takes the fuzzy output sets and transforms them back into numbers that indicate what action should take place, or decision should be made. Our simple fuzzy air conditioner is governed by two basic rules (real air conditioners would probably be governed by more than two rules), which use the two fuzzy input sets cold and hot. These fuzzy sets describe the temperature. The rules are associated with two fuzzy output sets, high and off which describe the settings for the air conditioner [5].

V. TYPE-2 FUZZY SYSTEMS

In this section, Type-2 fuzzy systems are presented. The structure of the Type-2 fuzzy rules is the same as for the Type-1 case because the distinction between Type-2 and Type-1 is associated with the nature of the membership functions [6]. Hence, the only difference is that now some or all the fuzzy sets involved in the rules are of Type-2. In a Type-1 fuzzy system, where the output sets are Type-1 fuzzy sets, we perform defuzzification in order to get a number, which is in some sense a crisp (type-0) representative of the combined output sets. In the Type-2 case, the output sets are of Type-2, so we have to use extended versions of Type-1 defuzzification methods [6]. The structure of Type-2 fuzzy logic system is shown below in Fig. 3.

A Type-2 membership grade can be any sub-set in [0, 1] the primary membership and corresponding to each primary membership, there is a secondary membership (which can also be in [0, 1]) that defines the possibilities for the primary membership. A Type-2 FLS is characterized by IF-THEN rules, where their antecedent or consequent sets are now of Type-2. Type-2 FLSs, can be used when the circumstances are too unknown to determine exact membership grade such as when the training data is affected by noise.

1. Fuzzifier: The fuzzifier maps a numeric vector x = $(x_1 \ldots . x_p)^{T} \in X_1 * X_2 * \ldots\ldots\ldots\ldots * X_p \equiv X$ into a Type-2 fuzzy into a Type-2 fuzzy set $\tilde{A}_x$ in X [9], an interval Type-2 fuzzy set in this case. We use Type-2 singleton fuzzifier, in a single on fuzzification, the input fuzzy set has only a one point on non zero membership [10].
2. Rules: The structure of rules in a Type-1 FLS and a Type-2 FLS is the same, but in the latter the antecedents and the consequents is represented by Type-2 fuzzy sets [10].
3. Type reducer: The type-reducer generates a Type-1 fuzzy set output, which is then converted in a numeric output through running the defuzzifier. This Type-1fuzzyset is also an interval set, for the case of our FLS we used center of sets (cos) type reduction [10] [8].
4. Defuzzifier: From the type- reducer, we obtain an interval set Y cos, to defuzzify it we use the average of yl and yr, so the defuzzified output of an interval singleton Type-2 FLS is Y(x) = $(y_{(1+yr)})/2$.

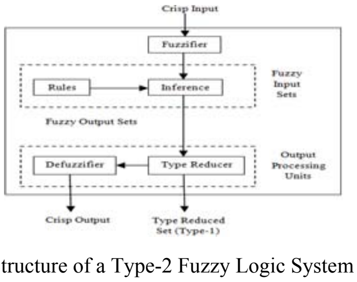


Fig. 3 Structure of a Type-2 Fuzzy Logic System [7] [8]

VI. Application of Type-2 Fuzzy Logic in Various Fields

In this section a representative account of the most successful applications of Type-2 fuzzy logic in various fields is presented. Type-2 fuzzy logic has been used to allow handling higher levels of uncertainty in real world complex problems. In the applications presented in this section the superiority of Type-2 over Type-1 fuzzy logic has been shown to be significant. The applications considered in these papers are diverse, ranging from medicine to social sciences, which show the importance of the use of Type-2 fuzzy logic for this kind of problems.

In Rubio Solis, A. and Panoutsos, G. [12], an interval Type-2 radial basis function neural network (IT2-RBF-NN) is proposed as a new modeling framework. In this functional equivalence of radial basis function neural networks (RBFNN) to a class of Type-1 fuzzy logic systems (T1-FLS) to propose a new interval Type-2 equivalent system, it is systematically shown that the type equivalence (between RBF and FLS) of the new modeling structure is maintained in the case of the IT2 system. A very good computational efficiency is demonstrated as a result of the systematic and automatic creation of IT2 linguistic information and the FOU.

In the proposed approach of Melin, P., Gonzalez, C., Castro.J, Mendoza O. and Castillo O. [13], an edge detection method based on the morphological gradient technique and generalized Type-2 fuzzy logic is proposed. The theory of alpha planes is used to implement generalized Type-2 fuzzy logic for edge detection. For the defuzzification process, the heights and approximation methods are used.

In the approach of Aisbett, J. and Rickard, J.T., [14], Centroids are practically important in Type-1 and Type-2 fuzzy logic systems as a method of defuzzification and type reduction is proposed. However, computational problems arise when membership functions (MF) have singleton spikes.

In the approach of Naim, S. and Hagras, H. [15], A general Type-2 fuzzy logic based approach for Multi-Criteria Group Decision Making," Fuzzy Systems (FUZZ), is purposed. Multi- Criteria Group Decision Making (MCGDM) is used for viewing decision making. MCGDM is a decision tool which it is used to find a unique agreement from number of decision makers and users by evaluating the unknown judgment among them. Several fuzzy logic based approaches have been used in MCGDM to handle the linguistic uncertainties and hesitancy.

In the proposed approach of Khanesar, M.A., Kayacan, E.Kaynak, O. and Saeys, W. [16], Sliding mode Type-2 fuzzy control of robotic arm using ellipsoidal membership functions is proposed. Several papers state that the performance of the Type-2 fuzzy logic systems is superior over their Type-1 counterparts, especially under noisy conditions. To show the effectiveness of the noise reduction capabilities of the Type-2 fuzzy logic systems, a novel Type-2 fuzzy membership function, ellipsoidal membership function, has recently been proposed.

In the proposed approach of Nguyen T., Khosravi A. Nahavandi S and Creighton D. [17], neural network and interval Type-2 fuzzy system for stock price forecasting is proposed. Stock price forecast has long been received special attention of investors and financial institutions. As stock prices are changeable over time and increasingly uncertain in modern financial markets, their forecasting becomes more important than ever before. An interval Type-2 fuzzy logic system (IT2 FLS) is employed as the second component of the hybrid forecasting method. The IT2 FLS's parameters are initialized through deployment of the k-means clustering method and they are adjusted by the genetic algorithm.

In the proposed approach of Pulido, M., Melin, P. and Castillo, O. [18], Optimization of ensemble neural networks with Type-2 fuzzy response integration for predicting the Mackey-Glass time series is proposed. The optimization of an ensemble neural network with fuzzy integration of responses based on Type-1 and Type-2 fuzzy logic is explained. Genetic algorithms are used as a method of optimization for the ensemble model in this case of study. The time series that is being considered is the Mackey-Glass benchmark.

In the proposed approach of Wati, D.A.R. and Jayanti, P.N. [19], Interval Type-2 Fuzzy Logic Controller of heat exchanger systems," Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME) is proposed. Heat exchanger systems are widely used in chemical plants. They are affected on load and disturbance of the process plant, change in operation condition and nonlinearity. In this paper, we studied the design of an interval Type-2 Fuzzy Logic Controller (FLC) for stirred tank heat exchanger systems. It controls the coolant flow in order to result in the desired temperature of output fluid.

In the proposed approach of Liang Zhao; Yanzhen Li; Yanjun Li [20], Computing with words for discrete general Type-2 fuzzy sets based on plane," Vehicular Electronics and Safety (ICVES), is proposed. General Type-2 fuzzy set (GT2FS) is the generalization of its Type-1 counterpart, which can better describe the nature of uncertainty. This paper presents computing with words (CWWs) for the discrete GT2FS based on plane representation. Firstly, it is introduced for the GT2FS. Secondly, we studied the numerical procedure of CWWs. Numerical examples are applied to assess the algorithm.

In the proposed approach of Farooq, U.and Gu, J. [21], A simple interval Type-2 fuzzy gain scheduling controller is designed for the stabilization and reference tracking of ball and plate system. The controller employs plate angles as the premise variables for gain scheduling and its stability is guaranteed through a set of linear matrix inequalities. MATLAB simulations are performed to validate the proposed controller where it is also compared with pole placement and Type-1 fuzzy logic controllers. It is shown that the proposed controller has better response and disturbance rejection capability and is robust to measurement noise and errors.

In the proposed approach of Soto, J.; Melin, P. and Castillo, O. [22], Optimization of Interval Type-2 and Type-1 Fuzzy Integrators in Ensembles of ANFIS Models with Genetic Algorithms is proposed. An optimization of interval Type-2 and Type-1 fuzzy integrators in ensembles of ANFIS models with genetic algorithms (GAs) is represented, in this with emphasis on its application to the prediction of chaotic time series, where the goal is to reduce the prediction error. The time series that was considered is the Mackey-Glass to test the experiments.

In the proposed approach of Gonzalez, C.I [23], A new approach based on generalized Type-2 fuzzy logic for edge detection is proposed. An edge detection method based on morphological gradient technique and generalized Type-2 fuzzy logic. The theory of alpha planes is used to implement generalized Type-2 fuzzy logic. For the test we used the method of defuzzification by height and approximation. The simulation results were obtained with a Type-1 fuzzy inference system (T1FIS), an interval Type-2 fuzzy inference system (IT2FIS) and with a generalized Type-2 fuzzy logic (GT2FIS). The proposed Type-2 fuzzy edge detection method was tested with benchmark images and synthetic images.

In the proposed approach of Yunrui Bi and Srinivasan, D. [24], Single intersection signal control based on Type-2 fuzzy logic is proposed. Single intersection is generally regarded as the elementary unit for solving traffic problem. A Type-2 fuzzy logic controller (T2FLC) for single intersection signal control is presented in this paper as Type-2 fuzzy logic can handle the imprecision, uncertainties and vagueness lying in the dynamic process more efficiently.

In the proposed approach of Farooq, U [25], Design and comparison of Type-1 and interval Type-2 fuzzy gain scheduling controllers for ball and beam system is proposed. The paper presents the design and comparison of Type-1 and interval Type-2 fuzzy logic controllers for ball and beam system.

In the proposed approach of Kumbasar, T [26], A Type-2 Fuzzy Cascade Control Architecture for Mobile Robots," Systems, Man, and Cybernetics (SMC), is proposed. The real-time path tracking control of mobile robots attracted considerable research interest since they inherit non holonomic properties and uncertainties caused by the internal dynamics or feedback sensors. In this paper, we studied a cascade control architecture, which includes the inner and outer control loops, for the path tracking control of mobile robots.

In the proposed approach of M. A. Sanchez, J. R. Castro, F. Perez-Ornelas, and O. Castillo, O [27], A hybrid method for IT2 TSK formation based on the principle of justifiable granularity and PSO for spread optimization is proposed. A new hybrid method for forming interval type 2 fuzzy inference systems (IT2 FIS) is shown. This methodology builds upon an existing Type-1 fuzzy inference system (T1 FIS) or from the output centers from any clustering algorithm, calculating the footprint of uncertainty (FOU) based on the implementation of the principle of justifiable granularity, and finally a particle swarm optimization algorithm (PSO) optimizes the spreads from First Order Takagi-Sugeno-Kang (TSK) type consequents to improve the coverage of the FOU. Focusing mainly in the coverage of the FOU, two datasets are used to demonstrate the effectiveness of FOU coverage in environments with noise, especially when the noise is on the outputs. These two datasets are a simple Fifth Order curve, and the iris benchmark dataset.

VII. Conclusion

In this paper a representative and concise review of Type-2 Fuzzy Logic of various applications was presented. Now's days Fuzzy Logic has got famous for using in various applications as discussed above, due to its ability to handle complexity. In this paper, a representative review of the most recent application of Type-2 Fuzzy Logic was given. Fuzzy Logic is gaining popularity due to handling of various uncertainties in the various fields. Fuzzy logic is also helpful to gain the data from the object knowledge and to predict the next result on the bases of previous knowledge.

References

[1] J. Sanders and E. Kandrot, CUDA by Example: An Introduction to General-Purpose GPU Programming. Addison-Wesley Professional, 2010.

[2] Owens, J.D.; Houston, M.; Luebke, D.; Green, S.; Stone, J.E.; Phillips, J.C., "GPU Computing," Proceedings of the IEEE, Vol. 96, No. 5, pp. 879, 899, May 2008 doi: 10.1109/JPROC.2008.917757

[3] R.H. Luke III, D. Anderson, J.M. Keller, and S. Coupland, "Fuzzy logic-based image processing using graphics processor units." In IFSA/EUSFLAT Conference, 2009, pp. 288–293.

[4] Mendel, Jerry M. "Fuzzy logic systems for engineering: a tutorial." Proceedings of the IEEE 83.3 (1995): 345–377.

[5] M.A. Martin and J.M. Mendel, "Flirtation: A Very Fuzzy Prospect: A Flirtation Advisor," Journal of Popular Cult., XI (1), pp. 1–41, 1995.

[6] N.N. Karnik, J.M. Mendel, and Q. Liang, "Type-2 Fuzzy Logic Systems,", IEEE Transactions on Fuzzy Systems, Vol. 7, No. 6, pp. 643–658, 1999.

[7] A. Khosravi, S. Nahavandi, D. Creighton, and D. Srinivasan, "Interval Type-2 Fuzzy Logic Systems for Load Forecasting: A Comparative Study," IEEE Transactions on Power Systems, Vol. 27, No. 3, pp. 1274–1282, 2012.

[8] E.A. Jammeh, M. Fleury, C. Wagner, H. Hagras, and M. Ghanbari, “Interval Type-2 Fuzzy Logic Congestion Control for Video Streaming Across IP Networks,” IEEE Transactions on Fuzzy Systems, Vol. 17, No. 5, pp. 1123–1142, 2009.

[9] C.F. Juang, R.B. Huang, and Y.Y. Lin, “A Recurrent Self-evolving Interval Type-2 Fuzzy Neural Network for Dynamic System Processing,” IEEE Transactions on Fuzzy Systems, Vol. 17, No. 5, pp. 1092–1105, 2009.

[10] P. Melin and O. Castillo, “A Review on Type-2 Fuzzy Logic Applications in Clustering, Classification and Pattern Recognition,” Applied Soft Computing, Vol. 21, pp. 568–577, 2014.

[11] C.H. Lee, F.Y. Chang, and C.M. Lin, “An Efficient Interval Type-2 Fuzzy CMAC for Chaos Time-Series Prediction and Synchronization,” IEEE Transactions on Cybernetics, Vol. 44, No. 3, pp. 329–341, 2014.

[12] Rubio Solis, A; Panoutsos, G., "Interval Type-2 Radial Basis Function Neural Network: A Modelling Framework," Fuzzy Systems, IEEE Transactions on, Vol. PP, No. 99, pp. 1,1 doi: 10.1109/TFUZZ.2014.2315656.

[13] Melin, P.; Gonzalez, C.; Castro, J.; Mendoza, O.; Castillo, O., "Edge Detection Method for Image Processing Based on Generalized Type-2 Fuzzy Logic," Fuzzy Systems, IEEE Transactions on, Vol. PP, No. 99, pp. 1,1 doi: 10.1109/TFUZZ.2013.2297159.

[14] Aisbett, J.; Rickard, J.T., "Centroids of Type-1 and Type-2 Fuzzy Sets When Membership Functions Have Spikes," Fuzzy Systems, IEEE Transactions on, Vol. 22, No. 3, pp. 685,692, June 2014 doi: 10.1109/TFUZZ.2014.2306973

[15] S. Naim and H. Hagras, “A General Type-2 Fuzzy Logic based Approach for Multi-criteria Group Decision Making,” in 2013 IEEE International Conference on Fuzzy Systems (FUZZ), 2013, pp. 1–8.

[16] M.A. Khanesar, E. Kayacan, O. Kaynak, and W. Saeys, “Sliding Mode Type-2 Fuzzy Control of Robotic Arm using Ellipsoidal Membership Functions,” in 2013 Asian Control Conference (ASCC), 2013, pp. 1–6.

[17] T. Nguyen, A. Khosravi, S. Nahavandi, and D. Creighton, “Neural Network and Interval Type-2 Fuzzy System for Stock Price Forecasting,” in 2013 IEEE International Conference on Fuzzy Systems (FUZZ), 2013, pp. 1–8.

[18] M. Pulido, P. Melin, and O. Castillo, “Optimization of Ensemble Neural Networks with Type-2 Fuzzy Response Integration for Predicting the Mackey-Glass Time Series,” in 2013 World Congress on Nature and Biologically Inspired Computing (NaBIC), 2013, pp. 16–21.

[19] D.A.R. Wati and P.N. Jayanti, “Interval Type-2 Fuzzy Logic Controller of Heat Exchanger Systems,” in 2013 3rd International Conference on Instrumentation, Communications, Information Tech. & Biomedical Engineering (ICICI-BME), 2013, pp. 141–146.

[20] L. Zhao, Y. Li, and Y. Li, “Computing with Words for Discrete General Type-2 Fuzzy Sets based on Alpha-Plane,” in 2013 IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2013, pp. 268–272.

[21] U. Farooq, J. Gu, and J. Luo, “An Interval Type-2 Fuzzy LQR Positioning Controller for Wheeled Mobile Robot,” in 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), 2013, pp. 2403–2407.

[22] Soto, J.; Melin, P.; Castillo, O., "Optimization of interval type-2 and type-1 fuzzy integrators in ensembles of ANFIS models with Genetic Algorithms," Nature and Biologically Inspired Computing (NaBIC), 2013 World Congress on, Vol., No., pp. 41,46, 12-14 Aug. 2013 doi: 10.1109/NaBIC.2013.6617876

[23] C.I. Gonzalez, J.R. Castro, G.E. Martinez, P. Melin, and O. Castillo, “A New Approach based on Generalized Type-2 Fuzzy Logic for Edge Detection,” in 2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013, pp. 424–429.

[24] Y. Bi, D. Srinivasan, X. Lu, and Z. Sun, “Single Intersection Signal Control based on Type-2 Fuzzy Logic,” in 2013 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS), 2013, pp. 25–31.

[25] Farooq, U.; Gu, J., "Design and comparison of type-1 and interval type-2 fuzzy gain scheduling controllers for ball and beam system," Information & Communication Technologies (ICICT), 2013 5th International Conference on, Vol., No., pp. 1,7, 14-15 Dec. 2013doi: 10.1109/ICICT.2013.6732788.

[26] T. Kumbasar and H. Hagras, “A Type-2 Fuzzy Cascade Control Architecture for Mobile Robots,” in 2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2013, pp. 3226–3231.

[27] M.A. Sanchez, J.R. Castro, F. Perez-Ornelas, and O. Castillo, “A Hybrid Method for IT2 TSK Formation based on the Principle of Justifiable Granularity and PSO for Spread Optimization,” in 2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013, pp. 1268–1273.

Converting Waste Agricultural Biomass into Electrical Energy – Indian Perspective

Navneet Kaur
Department of Electrical Engineering,
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab, India
E-mail: navneetkular@rediffmail.com

***Abstract*—India is one of the developing countries whose economy is largely based on agriculture. It constitutes the backbone of rural India whose inhabitants are more than 70% of total population. As a result, lots of agricultural wastes are generated and remain unutilized. Globally, 140 billion metric tons of biomass is generated every year from agriculture. This level of biomass can be converted to a massive amount of energy and raw materials. Equivalent to approximately 50 billion tons of oil, agricultural biomass waste transformed to energy can significantly displace fossil fuel, decrease emissions of greenhouse gases and supply renewable energy to some 1.6 billion people in developing countries, which still lack access to electricity. As raw materials, biomass wastes have striking potentials for large-scale industry and community-level enterprises.**

Keywords: Waste Agricultural Biomass, Renewable Energy, Electric Generation, Cost Effective

I. Introduction

Biomass takes the form of residual stalks, straw, leaves, roots, husk, nut or seed shells, waste wood and animal husbandry waste. Widely available, renewable, and almost free, waste biomass is an imperative resource. Moreover, there is always an increasing apprehension due to rapid exhaustion of fossil fuel resources for power generation and resultant pollution of the environment. Therefore, biomass is being considered as one of the substitute sources of electricity generation [1]. Rotten waste agricultural biomass emits methane and open burning by the farmers to clear the lands generates CO_2 and other local pollutants. Hence improper organization of waste agricultural biomass is contributing towards climate change, water and soil contamination, and local air pollution.

Besides, this waste is of high significance with respect to material and energy recovery. Biomass obtains its energy from the sun while the plants are growing. Plants transform solar energy into chemical energy during the process of photosynthesis. This energy is released as heat energy when the plant material is burned [2].

II. Conversion

In biomass power plants biomass fuel is burnt in boilers. The heat released from this process is used to heat and convert water into steam to turn a steam turbine which runs a generator to create electricity. Biomass is sometimes burned in combination with coal in boilers at power plants [3]. This process, called co-firing, is typically used to lessen air emissions and other ecological impacts from burning coal.

Biomass power plants release nitrogen oxides and a small amount of sulphur dioxide. The amounts emitted depend on the kind of biomass that is burned and the type of generator used. Although the burning of biomass also produces CO_2, the major greenhouse gas, it is considered to be part of the natural carbon cycle of the earth. The plants take up carbon dioxide from the air while they are growing and then return it to the air when they are burned, thereby causing no net increase. Biomass contains much less sulphur and nitrogen than coal; therefore, when biomass is co-fired with coal, sulphur dioxide and nitrogen oxides emissions are lower than when coal is burned alone. When the role of renewable biomass in the carbon cycle is considered, the carbon dioxide emissions that result from co-firing biomass with coal are lower than those from burning coal alone [4].

III. Modern Biomass Technologies in India

A. Heat and Steam from Sugarcane Leaf and Bagasse

Certain critical engineering design norms of the gasification system were first developed on a laboratory-scale model and then used to design a full-fledged commercial scale system with a thermal output of 1080 MJ/ h. This system comprises of a reactor, a gas conditioning system, a biomass feeding system and the instrumentation and controls [5].

1. ***Reactor:*** This was a downdraft, throatless and open-top reactor with an internal diameter of 75 cm and an active bed height of 1.25 m. It was designed for a heavy-duty cycle of 7500 hour per year operation. High temperature resisting firebricks conforming to IS 8 grade were used for the hot face followed by cold face insulation.
2. ***Gas conditioning system:*** A completely dry dust collection system eliminated altogether the problem of wastewater. This consisted of a high temperature char/ash coarse settler and a high efficiency cyclone separator. A specifically designed high temperature resisting induced-draft fan 3 ensured that the entire system is under negative pressure so that in the event of leaks, outside air got sucked into the system, but the combustible gas did not leak out. Thus, this design is very environment-friendly. The char-ash from the coarse settler and the cyclone was collected in barrels and emptied in an ash pit once every forty-five minutes. This char-ash which typically has a gross calorific value of 18.9 MJ/ kg can be briquetted to form an excellent fuel.

3. ***Biomass feeding system:*** This consisted of a scraper drag-out conveyor and a hopper to convey the biomass fuel from the storage pile to the reactor. The conveyor was completely enclosed.
4. ***Instrumentation and Control System:*** A Programmable Logic Controller (PLC)—based control system was designed to take automatic corrective actions under certain critical conditions. Thus, the biomass feeding and ash removal rates were fully controlled by this system. Besides, it also helped the operator in trouble-shooting by monitoring temperatures at various critical points in the gasification system. Automatic burner sequence controllers were provided for ignition of the producer gas.

The gasification system was extremely simple to operate. A cold start took about ten-fifteen minutes whereas a hot start was established in less than five minutes. Only two operators per shift of eight hours were required to operate the system, including the fuel and ash handling operations. The gasification system was successfully tested on sugarcane leaves and bagasse, sweet sorghum stalks and bagasse, bajra stalks etc.

B. Rice-Husk Based Gasifier

Biomass gasifiers capable of producing power from a few KW up to 1 MW capacity have been successfully developed indigenously. Indigenously developed small biomass gasifiers have successfully undergone stringent testing abroad. Biomass Gasifiers are now being exported not only to developing countries of Asia and Latin America, but also to Europe and USA. A large number of installations for providing power to small-scale industries and for electrification of a village or group of villages have been undertaken. The Biomass Gasifier Programme has been re-casted to bring about better quality and cost effectiveness. The programmes on biomass briquetting and biomass production are being reviewed and a new programme on power production linked to energy plantations on waste lands is proposed to be developed [5].

Installations: A total capacity of 55.105 MW has so far been installed, mainly for stand-alone applications.

1. A 5 x 100 KW biomass gasifier installation on Gosaba Island in Sunderbans area of West Bengal is being successfully run on a commercial basis to provide electricity to the inhabitants of the Island through a local grid.
2. A 500 KW grid interactive biomass gasifier, linked to an energy plantation, has been commissioned under a demonstration project.
3. A 4 X 250 kW (1.00 MW) Biomass Gasifier based project has recently been commissioned at Khtrichera, Tripura for village electrification.

C. MXP 10 Technology

These fuel fired boilers are offered in capacities upto 25 tons of steaming per hour. These boilers are offered in Shell and tube type construction, water tube construction and composite design. Thermax Solid fuel fired boilers are available designed to fire a wide range of solid fuels including agrowastes and other solid waste. These boilers are known for their ruggedness, and high efficiency, even when firing difficult fuels [5].

Process Description: The technology is a smoke tube; single pass boiler which can be fired on a number of fuels e.g. coal, husk, bagasse and wood. Depending on the type of fuel it has either a balanced draught or an induced draught system. The boiler consists mainly of three parts:

1. Refractory lined external furnace
2. Shell and tube exchanger pressure part.
3. Atmospheric water preheater (optional)

The furnace has a step grate or a fixed grate depending on the fuel e.g. for husk it has a step grate whereas for coal it has fixed grate. The fuel is charged through charging door to keep the fire going. The furnace is lined with refractory and insulation bricks. Furnace and steam generating section are placed in line. The draught is made available through the openings under the grate. In the case of coal firing an FD fan provides air for combustion. With the combination of FD and ID fan, balanced draught is maintained. Ash doors are provided to remove the ash from beneath the grate.

D. Biomass Gasifier

The biomass gasifier is essentially a chemical reactor where various complex physical and chemical processes take place. Four distinct processes take place in a gasifier, namely drying of the fuel (woody biomass), pyrolysis, combustion and reduction. Biomass is fed into gasifier at regular intervals. The equipment is designed in such a way that it takes air in controlled quantities, resulting in partial oxidation of biomass into producer gas. One Kg of biomass gets converted into 2.5 to 3.0 Nm^3 of gas with a calorific value of 1000-1300 Kcal per Nm^3, which would have the following composition CO- 15-20% CH_4- 1-4% CO_2-8-5%, N_2 –45-55%. The gas coming out of the gasifier is hot (200–2500C) and contains some contaminants, particulates and volatiles, which needs to be cooled and cleaned before feeding into the generators. The cooling cleaning system consists of scrubbers and associated accessories. Cold clean gas produced is fed to engine along with air. Woody biomass available in plenty is collected and transported to the gasifier through a skip charger [5].

Environmental Considerations: It converts a traditional low quantity fuel inconvenient for use into high quantity, combustible gaseous fuel with associated convenience. Such conversions are relatively at high efficiencies and result in total convenience and process control. Almost all environment pollution associated with biomass use can be eliminated. Initial investment and also the cost of generation of energy are just about the lowest amongst all known alternatives. Gestation period is very less and can be completed within 3 to 6 months.

A 10-15 MW agri-waste based power project has been set up jointly by **Punjab Biomass Power, Bermaco Energy, Archean Granites and Gammon Infrastructure projects Limited in Punjab.** The project uses locally available agricultural waste such as rice straw and sugar cane trash for fuel [6]. The total annual fuel requirement is around 120,000 tonnes of biomass, all of which will be acquired locally. Punjab has about 20 million acres under paddy yielding 100 million tonnes of rice straw. As rice straw is a poor fodder and fuel, farmers burn it in the fields and make way for the Rabi wheat crop. But now these wastes are being used for generating electricity. The project is expected to provide additional income to 15,000 farmers from the sale of agri waste. The project will be a major landmark in environment protection-converting agricultural waste to energy. Secondly, it will trim down the release of smoke and other pollutants caused by burning 100 million tonnes of wastes which could be used for earning carbon credits.

IV. Advantages

Sustainable production and utilization of biomass in electrical power generation can also solve the problems of rural unemployment, utilization of wasteland, and transmission losses in grid network [7]. Therefore, the system of biomass-based power generation is being given precedence in most of the developing nations including India. Unlike other renewable substances biomass supplies, pre-dried up to about 15% moisture content, can be stored for a much longer period of time without any complications. Besides electricity supply to the national grids, biomass offers great opportunities for decentralized power generation in rural areas at or near the points of use and thus can make villagers/small industries self-dependent in respect of their power requirements. It is necessary to find out the various properties like calorific value, chemical composition, reactivity towards oxygen, bulk density, etc to exploit biomass species in electricity generation.

V. Future of Biomass Energy in India

Future of biomass energy depends on providing consistent energy services at competitive cost. In India, this will happen only if biomass energy services can compete on a fair market. Policy priorities should be to orient biomass energy services towards market and to reform the market towards fair competition. Most cost-effective option is exploitation of waste materials. Potential availability of agro residues and wood processing waste in India can sustain 10,000 MW power [8]. Constant supply of biomass shall require production of energy crops (e.g., wood fuel plantations, sugar cane as feedstock for ethanol) and wood plantations for meeting growing non-energy needs. Land supply, enhanced biomass productivity, economic operations of plantations and logistics infrastructure are significant areas which shall determine future of biomass in India.

VI. Conclusion

This paper emphasized on possible use of waste agricultural biomass resources currently available in India for the production of electrical energy. There is a variety of biomass resources existing in the country and there is also huge opportunity for their conversion to various types of biofuels using different biomass conversion technologies. The availability of different types of agricultural crop residues, forest residues and wood processing waste makes them possible biofuel feedstocks. Also, the organic part of municipal solid waste, together with animal manure could play a key role as potential cellulosic feedstocks for the production of biogas.

The adoption of biomass as fuel for electricity generation in India can ease the financial strain relating to the heavy burden of fossil fuel costs and also develop local source of revenue for the people within the production chains. With the very high potential for biofuel production, the Government as well as private investors should take steps towards investing in agriculture for the production of energy crops and establishment of processing units for agricultural residues.

References

[1] Donald L. Klass, *An Introduction to Biomass Energy a Renewable Resource*, 1998, (www.bera1.org).

[2] Nisha Sriram and Mohammad Shahidehpour, *Renewable Biomass Energy*, 2005.

[3] P.R. Shukla, *Biomass Energy in India: Transition from Traditional to Modern*, The Social Engineer 6 (2), 1997.

[4] T.P.S Rajan, Eco-friendly fuels from perennial and renewable resources-Indian Scene, *Chemical Weekly*, 40 (11), 1995.

[5] UNEP, Converting Waste Agricultural Biomass into a Resource Compendium of Technologies, Report, *United Nations Environmental Programme Division of Technology*, Industry & Economics International Environmental Technology Centre Osaka/Shiga, Japan, 2009.

[6] N.H. Ravindranath and D.O. Hall, *Biomass Energy and Environment-A Developing Country Perspective from India*, Oxford University Press, Oxford, 1995.

[7] *Rural Energy in India*, www.indiasolar.com/bio.htm.

[8] P.R. Shukla, *Biomass Energy Future for India*, International Workshop, Rio De Janeiro, Brazil, 1997.

Speedup of Type-1 Fuzzy Logic Systems on Graphics Processing Units Using CUDA

Durlabh Chauhan[1], Satvir Singh[2], Sarabjeet Singh[3] and Vijay Kumar Banga[4]
[1,2]Department of Electronics & Communication Engineering,
SBS State Technical Campus, Ferozepur–152004, (Punjab) India
[3]Department of Computer Science & Engineering,
SBS State Technical Campus, Ferozepur–152004, (Punjab) India
[4]Department of Electronics & Communication Engineering,
Amritsar College of Engg. & Tech., Amritsar–143001, (Punjab) India
E-mail: [1]er.durlabh@gmail.com, [2]drsatvir.in@gmail.com,
[3]sarabjeet_singh13@yahoo.co, [4]v_banga@rediffmail.com

***Abstract*—Parallelcomputing is one of significant components of the High Performance Computing (HPC) and is being used to solve problems, which are large and complex in nature. Fuzzy Logic System (FLS) is a problem that becomes computationally intensive with increase in number of inputs and/or fuzzy rules. Running an FLS is highly parallel in nature, therefore, can be implemented in parallel on GPU using CUDA. In this paper, various fuzzy computations viz. rule firing, implication, aggregation and defuzzification are performed in parallel. Multiple threads are run at a time to fire multiple fuzzy rules, simultaneously, that reduces the overall FLS execution time. It is observed from simulation results that GPU works faster as compared to CPU when either number of inputs is increased or number of fuzzy rules.**

Keywords: High Performance Computing, Fuzzy Logic Systems, General Purpose Computing on Graphics Processing Unit, Compute Unified Design Architecture

I. Introduction

Artificial Intelligence (AI) undoubtedly is the backbone of the emerging technological advancements, however, implementation involves intensive computation owing to increased data sizes. Methods to improve the runtime performance construe the areas of research to reduce mathematical computations and parallelization of algorithm in hardware. Graphics Processor Units (GPU), plays a major role in gaming and graphics applications and allows for general purpose programming on remarkably fast parallel hardware using a Single Instruction Multiple Data (SIMD) programming architecture. Fuzzy Logic, introduced by Zadeh [1], [2], is one of the exigent part of AI and possess inherent parallel nature. General Purpose computing on GPU (GPGPU) is targeted by many researchers to speed up complicated algorithms especially, AI algorithms and their applications [3]. Ander-son *et al.* presented a GPU solution for the fuzzy C-means clustering algorithms [4]. Earlier this solution used OpenGL and Cg (graphics libraries) to achieve approximately two folds of computational speedup for some clustering profiles using NVIDIA 8800 GPU. They later generalized the system for the use of non-Euclidean metrics [5]. Further, Sejun Kim describes the method used to adapt a multilayer tree structure composed of fuzzy adaptive units into CUDA (Compute Unified Device Architecture) platforms [6].

In [7], Chiosa and Kolb present a framework for mesh clustering solely implemented on the GPU with a new generic multilevel clustering technique. Chia *et al.,* have proposed the implementation of a zero-order TSK-Fuzzy Neural Network (FNN) on GPUs to reduce training time in [8].

Harvey *et al.,* have presented a GPU solution for fuzzy inference system in [9]. Anderson *et al.,* present a parallel implementation of fuzzy inference on GPU using CUDA in[10]. Two folds of speed improvement of this naturally parallel algorithm have been achieved under typical inference profiles. One problem with this system and the implementation on GPU is that they both rely upon OpenGL and Cg libraries, which makes system generalization difficult for new comers to GPGPU. Further, Ngo *et al.,* report an implementation of Interval Type-2 FLS on GPU using CUDA with a tremendous speedup of 30 folds in [11].

In this paper, parallel functionality with CUDA programming model, for parallel implementation of a Type-1 Sugeno FLS on GPU, is investigated with increased number of inputs and fuzzy rules. After this brief historical background rest of this paper is outlined as follows: Section II targets CUDA programming model for GPGPU. Section III introduces Type 1 FLS along with the scope of parallelism wherever possible. Section IV presents simulation results and a speedup comparison of serial and parallel computing using CPU and GPU, respectively. Finally, section V concludes the research workand presents future off-shoots.

II. CUDA Programming Model

CUDA is a parallel computing platform and programming model introduced by NVIDIA that increases computing performance substantially by harnessing the power of parallelism of the GPU. CUDA gives program developers the direct access to the virtual instruction set and memory of the parallel computational elements in CUDA enabled GPUs. The CUDA platform is accessible to software developers through CUDA accelerated libraries, compiler directives (such as Open ACC), and extensions to industry-standard programming languages, including C, C++ and FORTRAN. C/C++ programmers use

CUDAC/C++, compiled with nvcc which is NVIDIA'S LLVM-based C/C++ compiler. A C/C++ program using CUDA can interface with one GPU or multiple GPUs and can be identified and utilized in parallel, allowing for unprecedented processing power on desktop computers.

CUDA allows multiple kernels to be run simultaneously on GPU cores. CUDA refers to each kernel as a grid. A grid is a collection of blocks. Each block runs the same kernel, however, is independent of each other (this has significance in terms of access to memory types). A block contains threads, which are the smallest divisible unit on a GPU.

A thread block is a number of SIMD threads that work on core at a given time. Threads can exchange information through the shared memory and can be synchronized. The operations are systematized as a grid of thread blocks. For parallel operation the programming model allows a developer to partition a program into several subprograms, each of which is executed independently on a block. Each subprogram can be further divided into finer pieces that perform the same function but execute on different threads independently within the same block. For data set parallelism, data sets can be divided in to smaller chunks that are stored in the shared memory, and each chunk is visible to all threads of the same block. This local data arrangement approach reduces the need to access off-chip global memory, which reduces data access time.

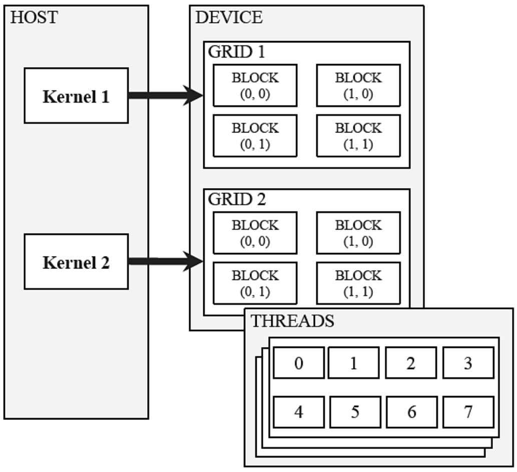


Fig. 1 CUDA Architecture

The next critical component of a CUDA application is the memory model. There are multiple types of memory and each has different access times. The GPU is broken up into read-write per thread registers, read-write per thread local memory, read-write per-block shared memory, read-write per-grid global memory, read-only per-grid constant memory, and read-only per-grid texture memory. Texture and constant memory have relatively small access latency times, while global memory has the largest access latency time. Applications should minimize the number of global memory reads and writes. This is typically achieved by having each thread read its data from global memory and store its content into shared memory.

The basic structure of a CUDA code comprises of allocation of memory space (using cuda Malloc function) on device (GPU) and (using regular malloc function) on host (CPU). Data which is copied from the host to the device for the call of kernel routine to be executed on the GPU (using function cuda Memcpy) also defines the number of threads and their physical structure. Kernel is prefixed with the global keyword. Results are transferred from GPU to CPU in the same fashion as data is copied from host to device.

III. Scope of Parallelism in Type-1 FLS

Theory of FLS given by Zadeh a fuzzy set is defined for a particular domain, and it is characterized by a membership function that maps elements from the domain to a real valued numbers [1], [2]. Mendel and many researchers gave numerous methods to design and implement FLS for various applications [12]–[15]. Theory of FLS given by Zadeh a fuzzy set is defined for a particular domain, and it is characterized by a membership function that maps elements from the domain to a real valued numbers.

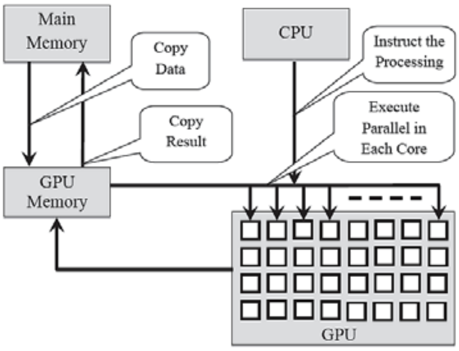


Fig. 2 CUDA Process Flow

In this paper every input has been fuzzified using four arbitrary located Gaussian membership functions that are characterized by two parameters, i.e., mean (*m*) and standard deviation (σ). Gaussian membership function is symmetrical about its mean and is expressed mathematically as (1)

$$\mu(x) = \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right) \tag{1}$$

In this paper, an FLS with 4 inputs and 1 output is subjectedto parallel computation for forecasting of Mackey-Glass timeseries [14]. For implication and aggregation, min and maxoperator are investigated. The defuzzification is performedwith the height defuzzification method as symmetrical shaped Gaussian fuzzy sets have been used to fuzzify consequent.Output of the height defuzzifier is given by (2)

$$y = \frac{\sum_{i=1}^{M} C_i \mu(x_i)}{\sum_{i=1}^{M} \mu(x_i)} \tag{2}$$

Here C represents the location of singleton consequents fuzzy sets and (x_i) represent clipping level for each rule after implication and M represents total number of fuzzy rules. However, for implementation of the FLS defined above in CUDA, the foremost and the most crucial step is to allocate the memory space for the data sets to be used in the system,as the choice and format of data affects performance of thealgorithm.

Scope of possible parallel computational processing is discussed as follows: Harvey *et al.,* [9] and Anderson *et al.,* [10] in their respective work have given a vast scope of parallelism for Type-1 FLS. In the same fashion Ngo *et al.,* [11] presented a novel scope of parallelism for Interval Type-2 FLS. However, in all these implementations the emphasis was laid to parallelize the number of fuzzy rules and discrete levels for a typical FLS with two inputs and single output. So, a single FLS was computed with parallel rule inference on GPU using CUDA. However, a typical FLS can be processed multiple times with fixed fuzzy rules and discrete levels but varying inputs. Here lies our scope of parallelism, we construe our code to compute multiple FLS in parallel on GPU as a FLS serially on CPU will consume more time.

IV. Parallel I Mplementation

Two M×N dimensional 'Mean' and 'Sigma' matrices are used in this implementation where M denotes number of fuzzy rules and N is number of antecedents. These matrices hold mean and sigma values of Gaussian membership function in accordance with fuzzy rules used in the FLS. An M×1 dimensional 'Consequent' matrix contains only mean values of the consequent fuzzy sets as systems uses height defuzzificationmethod given by equation (2). Multiple inputs are provided to the systems collectively in the form of an L×N dimensional input matrix, X, where L is the number of inputs.The CPU executes rule firing sequentially with a single input at a time and causes more computational time. Whereas, multiple threads are run at the same time to fire multiple rules simultaneously using system matrices, i.e., 'Mean', 'Sigma' and 'Consequent' matrices,which reduce the execution time for the FLS. A kernel function is initialized from CPU to pass inputs in parallel to various GPU cores. Copying system matrices everytime along with input vectors and increases the GPU processing time. Therefore, system matrices are copied only once and subsequently require only input vectors those are passed to GPU in parallel to enhance the GPU performance.

V. Results Discussion

The speed up performance with GPU implementation of a Type-1 Sugeno FLS was compared with that of CPU implementation. Intel Core 2 Duo system under experimentation has 2GB of system RAM, and Windows 7 platform. The GPU used here works on nVIDIAGeforce GTX 650 with 1024 MB of texture memory, 192 stream processors, and PCI Express X16.

The number of antecedents is fixed to 4 and consequentto 1, the number of rules is varied between 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 and inputs were varied as 128, 256, 512, and 1024. A set of input data is obtained from Mackey-Glass time series for experimentation. Ratio of CPU to GPU runtimes with respect to number of fuzzy rules has been tabulated in Table I and presented graphicallyin Fig. 3.

TABLE I Fuzzy Rules *vs.* CPU to GPU Speedup Ratio

Number of Fuzzy Rules	CPU to GPU Runtime Speedup Ratio			
	128 Inputs	256 Inputs	512 Inputs	1024 Inputs
10	1.937	4.875	5.034	5.488
20	2.437	3.032	4.548	6.132
30	1.340	2.319	4.319	6.544
40	2.000	2.476	3.968	7.063
50	1.516	3.532	4.410	7.063
60	1.730	3.000	4.365	7.185
70	1.602	3.205	4.509	7.134
80	1.500	2.742	4.500	7.832
90	1.709	2.577	4.446	5.488
100	1.624	2.504	4.652	6.132

Here, it can be observed clearly that advantage of GPU computing has increased with increased input data vector sizes. In another simulation results, it is observed that CPU to GPU speedup time improves with increase in fuzzy rules. Computational time on CPU varies significantly due to already always running applications at the back end.Therefore, to present fair comparison serial and parallel timing analysis experiments with same FLS have been repeated 30 times. Average of CPU to GPU speedups for 30 monte-carlosimulations during serial and parallel computations of rule firing, implication, aggregation and defuzzification have beenpresented in Fig. 4–7.

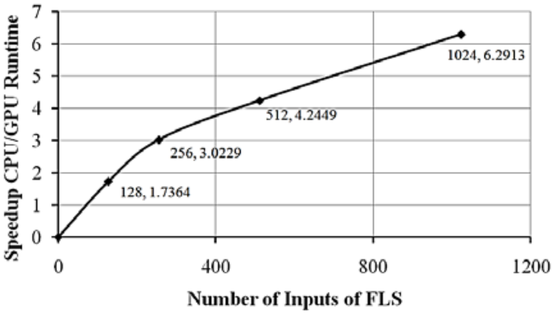


Fig. 3 CPU to GPU Speedup Ratio *vs* Inputs Data Length

The overall performance of CPU to GPU speedup timingswith respect to collective number of inputs, presented to the FLS, is shown in Fig. 8 that depicts that larger the number offuzzy rules better is the speedup performance.

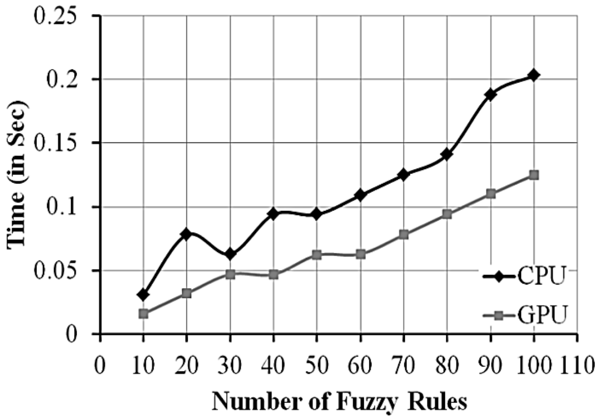


Fig. 4 CPU and GPU Runtime Comparison for 128 Inputs

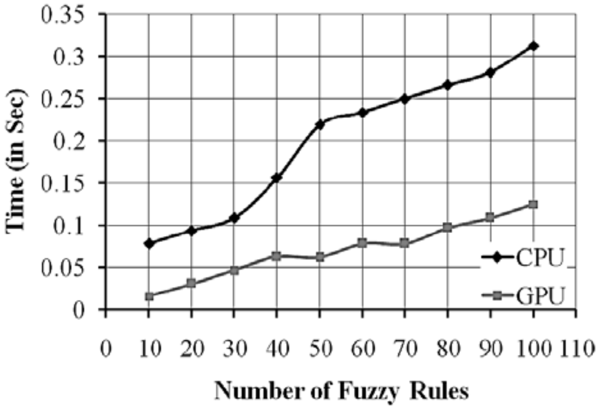


Fig. 5 CPU and GPU Runtime Comparison for 256 Inputs

VI. Conclusion and Future Work

This paper has demonstrated the implementation and comparative runtime performances of a typical Sugeno Type-1 FLSon a GPU and CPU without the use of a graphics API which is flexible, scalable, and can be used by any researcher with knowledge of C. It has been demonstrated that the CPU works equally fast as GPU when the system is small. As the number of rules or the number of inputs increase the GPU outperforms the CPU runtime. Here, in the work nearly 7.83 times speedup could be achieved as 1024 inputs supplied in parallel to the FLS ported on GPU. The GPU has an initial setup overhead of kernel loading and memory transfer, however, subsequent parallel computations leads to a small increase in processing time despite a substantial increase in computational load. On the other hand, CPU has no initial cost, but computation time grows linearly with computational load much beyond GPGPU runtime.

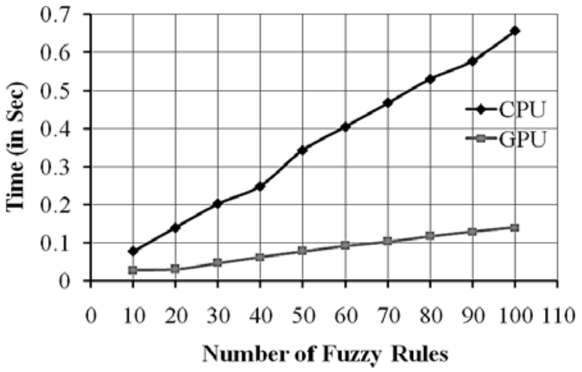


Fig. 6 CPU and GPU Runtime Comparison for 512 Inputs

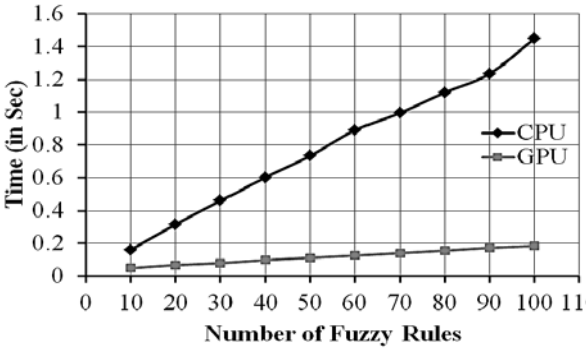


Fig. 7 CPU and GPU Runtime Comparison for 1024 Inputs

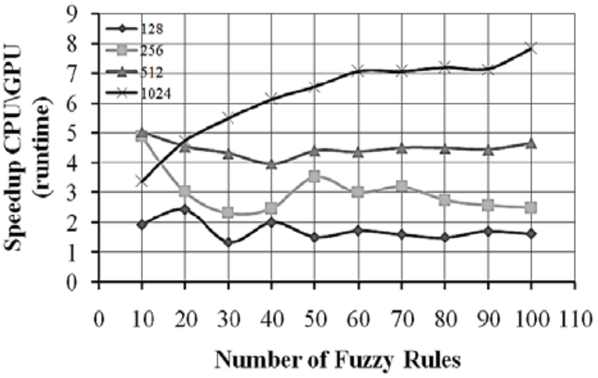


Fig. 8 CPU to GPU Speedup Ratio for Various Input Data Sets

The parallelization of more FLS applications is next on ouragenda. That will follow implementation of Interval Type-2 FLSs and Generalized Type-2 FLSs for various applications that require much more computational time otherwise.GPGPU is also possibly investigated to on Evolutionary Algorithms those are computational intensive and parallel innature.

References

[1] L.A. Zadeh, “Fuzzy Sets,” *Information and Control*, Vol. 8, No. 3, pp. 338–353, 1965.

[2] L.A. Zadeh, “Fuzzy Logic and Approximate Reasoning,” *Synthese*, Vol. 30, pp. 407–428, 1975.

[3] S. Singh, S. Singh, V.K. Banga, and D. Chauhan, “CUDA for GPGPU Applications-A Survey,” in *National Conference on Contem-poraryTechniques & Technologies in Electronics Engineering*, Murthal,Sonepat, India, March 2013, p. Accepted.

[4] D.T. Anderson, R.H. Luke, and J.M. Keller, “Speedup of FuzzyClustering through Stream Processing on Graphics Processing Units,” *IEEE Transactions on Fuzzy Systems*, Vol. 16, No. 4, pp. 1101–1106, 2008.

[5] D. Anderson, R.H. Luke, and J.M. Keller, “Incorporation of non-EuclideanDistance Metrics into Fuzzy Clustering on Graphics Processing Units,” in *Analysis and Design of Intelligent Systems using Soft Computing Techniques*. Springer, pp. 128–139, 2007.

[6] S. Kim and D. Wunsch, “A GPU based Parallel Hierarchical Fuzzy ART Clustering,” in *The 2011 International Joint Conference on Neural Networks (IJCNN)*, pp. 2778–2782, 2011.

[7] I. Chiosa and A. Kolb, “GPU-based Multilevel Clustering,” *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, No. 2, pp. 132–145, 2011.

[8] C.F. Juang, T.C. Chen, and W.Y. Cheng, “Speedup of Implementing Fuzzy Neural Networks with High-dimensional Inputs through Parallel Processing on Graphic Processing Units,” *IEEE Trans-actions on Fuzzy Systems*, Vol. 19, No. 4, pp. 717–728, 2011.

[9] N. Harvey, R. Luke, J.M. Keller, and D. Anderson, “Speedup of Fuzzy Logic through Stream Processing on Graphics Processing Units,”in 2008. *CEC IEEE World Congress on Computational Intelligence* and *Congress on Evolutionary Computation*, pp.3809–3815, 2008.

[10] D. Anderson and S. Coupland, “Parallelisation of Fuzzy Inference on a Graphics Processor Unit using the Compute Unified Design Architecture,”in *Proceedings of the UK Workshop on Computational Intelligence (UKCI’08)*, pp. 1–6, 2008.

[11] L.T. Ngo, D.D. Nguyen, C.M. Luong *et al.,*., “Speedup of Interval Type 2 Fuzzy Logic Systems based on GPU for Robot Navigation,” *Advances in Fuzzy Systems*, Vol. 2012, pp. 4, 2012.

[12] J.M. Mendel, “Fuzzy Logic Systems for Engineering: A Tutorial,”*Proceedings of the IEEE FUZZ*, Vol. 83, No. 3, pp. 345–377, 1995.

[13] G.C. Mouzouris and J.M. Mendel, “Non-singleton Fuzzy Logic Systems: Theory and Application,” *IEEE Transactions on Fuzzy Systems*, Vol. 5, No. 1, pp. 56–71, 1997.

[14] N.N. Karnik and J.M. Mendel, “Applications of Type-2 Fuzzy Logic Systems to Forecasting of Time-series,” *Information Sciences*, Vol. 120, No. 1, pp. 89–111, 1999.

[15] J.M. Mendel, “Uncertainty, Fuzzy Logic, and Signal Processing,” *Signal Processing*, Vol. 80, No. 6, pp. 913–933, 2000.

Author Index

International Conference on Communication, Computing and Systems 2014

SBS State Technical Campus and City Ferozepur

The Institute was established by the Government of Punjab in 1995 with the name Shaheed Bhagat Singh College of Engineering & Technology as a tribute to the great martyr Bhagat Singh. In the academic year 2011–12, it was upgraded to the status of a technical campus and rechristened as Shaheed Bhagat Singh State Technical Campus, Ferozepur. The Institute is fully promoted by the Punjab Government and is registered as a society under the Societies Registration Act 1860. Its affairs are administered by a Board of Governors. The Technical Campus, with its lush green state-of-the-artcampus spread about 100 acres is situated on Ferozepur–Moga Road (NH-95), about 4 km away from the Ferozepur city.

The City, Ferozepur is situated on the Indo-Pakistan International border. It is well connected by road and rail with important cities like Amritsar, Ludhiana, Jalandhar, Chandigarh, Delhi and rest of the country. As per the annals of history, Ferozeshah Tughlaq founded the city in the fourteenth century. However, it is also believed that its founder is Feroze Khan, one of the Bhatti chiefs. Having a rich heritage, it indeed has maintained its name. In the undivided India, it has been acentre of trade and commerce. Ferozepur, with numerous holy shrines, historical places and memorials in and around, has earned a rare status. The hallowed Samadhis of the martyrs: Bhagat Singh, Rajguru and Sukhdev, the Saragarhi Gurudwara commemorating the heroic sacrifice of Twenty One Sikh soldiers at Saragarhi post in Baluchistan and the Jain Swetambar Temple at Zira constructed in 1890 AD are some of the places of pilgrimage showing the spirit of universal brotherhood.

Supported & Sponsored by

ISTE — Indian Society for Technical Education

Optiwave

IETE — The Institution of Electronics and Telecommunication Engineers

सह वीर्य करवावहै

TEQIP

Group Excel India
www.groupexcelindia.com

Excel India Publishers

EXCEL INDIA PUBLISHERS
91 A, Ground Floor, Pratik Market, Munirka, New Delhi-110067
Call: +91-11-2671 1755/ 2755/ 3755/ 5755• Fax: 011-2671 6755
e-mail: publishing@groupexcelindia.com • Web: www.groupexcelindia.com

ISBN
9 789383 842681

₹ 1500 □ US$ 65