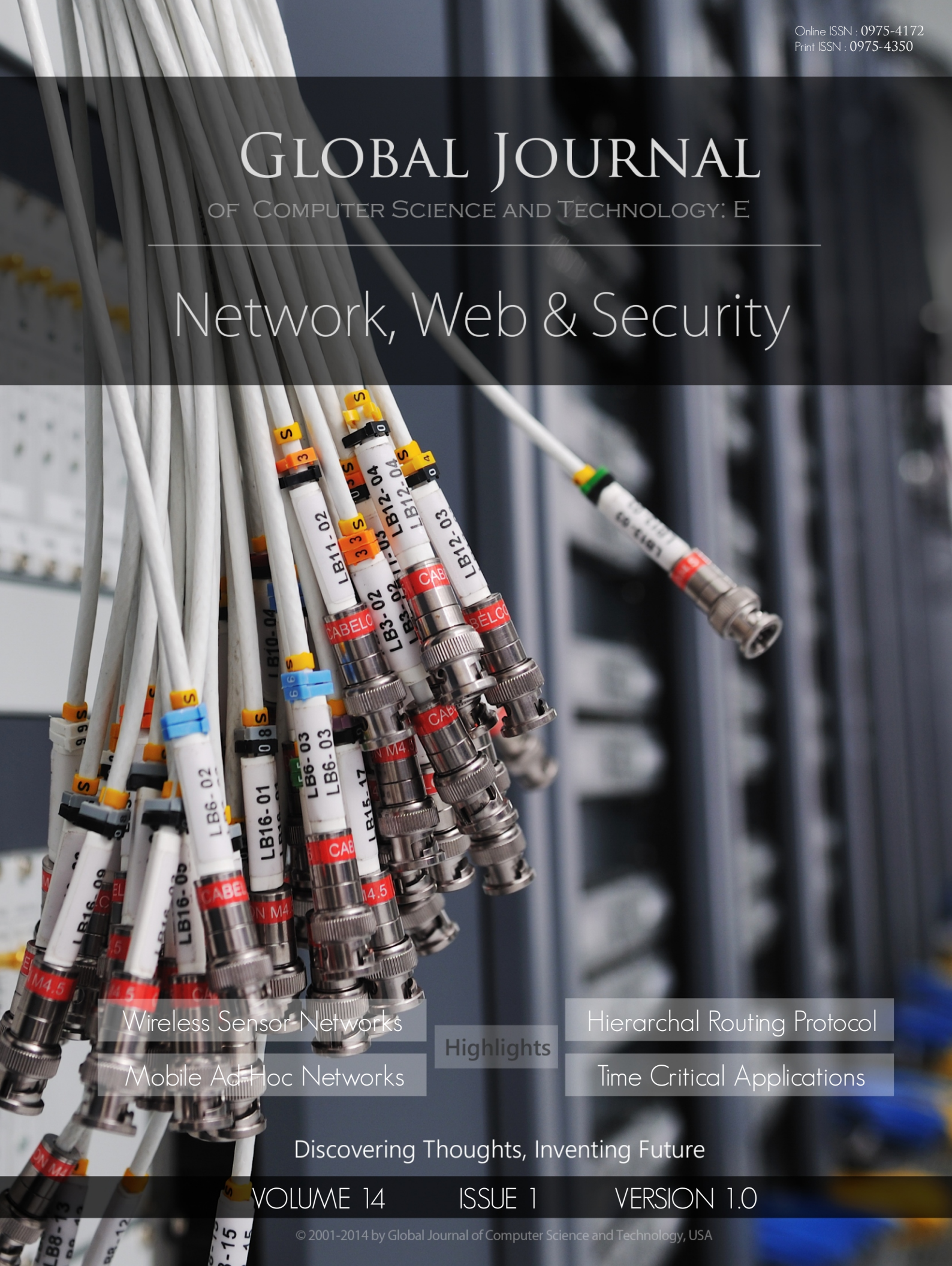
Online ISSN : 0975-4172
Print ISSN : 0975-4350

Global Journal

of Computer Science and Technology: E

Network, Web & Security



Highlights

- Wireless Sensor Networks
- Mobile Ad-Hoc Networks
- Hierarchal Routing Protocol
- Time Critical Applications

Discovering Thoughts, Inventing Future

VOLUME 14 ISSUE 1 VERSION 1.0

Global Journal of Computer Science and Technology: E
Network, Web & Security

Global Journal of Computer Science and Technology: E
Network, Web & Security

Volume 14 Issue 1 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

*(A Delaware USA Incorporation with "Good Standing"; **Reg. Number: 0423089**)*

Sponsors: *Open Association of Research Society*
Open Scientific Standards

Publisher's Headquarters office

Global Journals Headquarters
301st Edgewater Place Suite, 100 Edgewater Dr.-Pl,
Wakefield MASSACHUSETTS, Pin: 01880,
United States of America
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Incorporated
2nd, Lansdowne, Lansdowne Rd., Croydon-Surrey,
Pin: CR9 2ER, United Kingdom

Packaging & Continental Dispatching

Global Journals
E-3130 Sudama Nagar, Near Gopur Square,
Indore, M.P., Pin:452009, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investors@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:
22 USD (B/W) & 50 USD (Color)
Yearly Subscription (Personal & Institutional):
200 USD (B/W) & 250 USD (Color)

Integrated Editorial Board
(Computer Science, Engineering, Medical, Management, Natural Science, Social Science)

Dr. Bart Lambrecht
Director of Research in Accounting and FinanceProfessor of Finance
Lancaster University Management School
BA (Antwerp); MPhil, MA, PhD (Cambridge)

Dr. Carlos García Pont
Associate Professor of Marketing
IESE Business School, University of Navarra
Doctor of Philosophy (Management), Massachusetts Institute of Technology (MIT)
Master in Business Administration, IESE, University of Navarra
Degree in Industrial Engineering, Universitat Politècnica de Catalunya

Dr. Fotini Labropulu
Mathematics - Luther College
University of ReginaPh.D., M.Sc. in Mathematics
B.A. (Honors) in Mathematics
University of Windso

Dr. Lynn Lim
Reader in Business and Marketing
Roehampton University, London
BCom, PGDip, MBA (Distinction), PhD, FHEA

Dr. Mihaly Mezei
ASSOCIATE PROFESSOR
Department of Structural and Chemical Biology, Mount Sinai School of Medical Center
Ph.D., Etvs Lornd University
Postdoctoral Training,
New York University

Dr. Söhnke M. Bartram
Department of Accounting and FinanceLancaster University Management SchoolPh.D. (WHU Koblenz)
MBA/BBA (University of Saarbrücken)

Dr. Miguel Angel Ariño
Professor of Decision Sciences
IESE Business School
Barcelona, Spain (Universidad de Navarra)
CEIBS (China Europe International Business School).
Beijing, Shanghai and Shenzhen
Ph.D. in Mathematics
University of Barcelona
BA in Mathematics (Licenciatura)
University of Barcelona

Philip G. Moscoso
Technology and Operations Management
IESE Business School, University of Navarra
Ph.D in Industrial Engineering and Management, ETH Zurich
M.Sc. in Chemical Engineering, ETH Zurich

Dr. Sanjay Dixit, M.D.
Director, EP Laboratories, Philadelphia VA Medical Center
Cardiovascular Medicine - Cardiac Arrhythmia
Univ of Penn School of Medicine

Dr. Han-Xiang Deng
MD., Ph.D
Associate Professor and Research
Department Division of Neuromuscular Medicine
Davee Department of Neurology and Clinical NeuroscienceNorthwestern University
Feinberg School of Medicine

Dr. Pina C. Sanelli
Associate Professor of Public Health
Weill Cornell Medical College
Associate Attending Radiologist
NewYork-Presbyterian Hospital
MRI, MRA, CT, and CTA
Neuroradiology and Diagnostic Radiology
M.D., State University of New York at Buffalo,School of Medicine and Biomedical Sciences

Dr. Roberto Sanchez
Associate Professor
Department of Structural and Chemical Biology
Mount Sinai School of Medicine
Ph.D., The Rockefeller University

Dr. Wen-Yih Sun
Professor of Earth and Atmospheric SciencesPurdue University Director
National Center for Typhoon and Flooding Research, Taiwan
University Chair Professor
Department of Atmospheric Sciences, National Central University, Chung-Li, TaiwanUniversity Chair Professor
Institute of Environmental Engineering, National Chiao Tung University, Hsin-chu, Taiwan.Ph.D., MS The University of Chicago, Geophysical Sciences
BS National Taiwan University, Atmospheric Sciences
Associate Professor of Radiology

Dr. Michael R. Rudnick
M.D., FACP
Associate Professor of Medicine
Chief, Renal Electrolyte and Hypertension Division (PMC)
Penn Medicine, University of Pennsylvania
Presbyterian Medical Center, Philadelphia
Nephrology and Internal Medicine
Certified by the American Board of Internal Medicine

Dr. Bassey Benjamin Esu
B.Sc. Marketing; MBA Marketing; Ph.D Marketing
Lecturer, Department of Marketing, University of Calabar
Tourism Consultant, Cross River State Tourism Development Department
Co-ordinator , Sustainable Tourism Initiative, Calabar, Nigeria

Dr. Aziz M. Barbar, Ph.D.
IEEE Senior Member
Chairperson, Department of Computer Science
AUST - American University of Science & Technology
Alfred Naccash Avenue – Ashrafieh

Contents of the Volume

Global Journal of Computer Science and Technology: E Network, Web & Security
Volume 14 Issue 1 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Conference Rating

By Ankur Gupta
San Jose State University, India

Abstract- Conf Rate is a website designed to gather tweets regarding various conferences occurring around the corners of world and applying an algorithm to rate them on the sentimental analysis of tweets. When we talk of technolog stack implementation, Conf Rate uses Twython; a python based twitter API which connects to twitter on oauth (open authentication) and offers functionality like twitter search. The technique is to use the real time data from the twitter; commonly known as “tweets”. This information is categorized into positive or negative on the basis of the “tweet sentiments”. The Natural Language Toolkit (NLTK) is a powerful tool in Python which identifies tweet sentiments using Bayes classifier. The use of Django 1.3.1 web framework (a python framework) provides a lightweight web server with MVC framework which easily integrates with our python scripts. For designing user interface, html5 using jquery has been used. Finally, the application is deployed on the cloud using AWS (Amazon web services) Elastic Cloud Computer or commonly known as EC2, providing more profitability and reliable browsing and dependability from market perspective.

GJCST-E Classification : G.1.2

CONFERENCERATING

Strictly as per the compliance and regulations of:

Conference Rating

Ankur Gupta

Abstract- Conf Rate is a website designed to gather tweets regarding various conferences occurring around the corners of world and applying an algorithm to rate them on the sentimental analysis of tweets. When we talk of technolog stack implementation, Conf Rate uses Twython; a python based twitter API which connects to twitter on oauth (open authentication) and offers functionality like twitter search. The technique is to use the real time data from the twitter; commonly known as "tweets". This information is categorized into positive or negative on the basis of the "tweet sentiments". The Natural Language Toolkit (NLTK) is a powerful tool in Python which identifies tweet sentiments using Bayes classifier. The use of Django 1.3.1 web framework (a python framework) provides a lightweight web server with MVC framework which easily integrates with our python scripts. For designing user interface, html5 using jquery has been used. Finally, the application is deployed on the cloud using AWS (Amazon web services) Elastic Cloud Computer or commonly known as EC2, providing more profitability and reliable browsing and dependability from market perspective.

I. Introduction

Conferences happen on a day to day basis; some are good enough and some are not. As a novice in a particular industry, it becomes difficult to decide at times if it will be profitable enough to attend a particular conference or not? At times, it might happen that there is a tradeoff between the conferences you have to choose to attend. Consequently, nothing is gained by attending conferences that do not conform to their objectives. This offers the user a wide range of choices to choose from. To make a wise choice, the user needs to know which of those conferences are worth attending. This gives rise to the need for a tool that can fetch the reviews about the conferences from the top rated social networking sites such as Twitter. Thus, our goal is to design an application or a website that can assist people in getting conference rating along with visualizing the data gathered on the basis of tweets.

People or professionals attend conferences like IBM Information on demand, IBM Digital Experiences, Oracle World, Dream Force etc and provide their feedback on Twitter. Some may post their experience or provide a feedback about that conference. So the sole idea of our website is to offer users, the ratings of the conferences so that they are acquiesced with that particular conference in a much better way and hence plan their schedule accordingly. The user obtains the reviews in the form of rating such as Must Go, Go, Fair and Useless.

Author: *San Jose State University, Shastri Nagar, Delhi, India.*
e-mail: anky689@gmail.com

Conference with a very negative feedback will be rated as a useless conference.

Please note that we do not recommend any conference as a useless conference, it's our algorithm that suggests the viability of a particular conference not to be so good to attend only on the basis of twitter sentiment analysis. While at the same time, if a user search for a conference that does not exists, system may prompt an error message saying that you are looking for a conference that doesn't exists as no data is found for that particular occurrence. Added, if a user simply clicks on analyze button without entering any data, it redirects to a page saying error.

While the concept of Conf Rate is not entirely unique, the edge that has been added to it gives the entire concept a new and upgraded view as use of natural language toolkit with libraries such as stop words corpus, words frequency, classifier etc. that could easily be used to analyze the sentiments. Although, many work has been done so far around on twitter sentiment analysis but

When it comes to conference rating, there is no consolidated single tool or website to help you do so. Targeted users for this particular application are professionals, students, freelancers and people who are interested to attend to the conferences.

II. Motivation

Working marathon hours is a part of most of the professional's schedule and many technical conferences are organized on a day to day basis. Say for example, I am new to a particular industry with avid interest in automobile and want to know if an upcoming conference by General Motors would be worth attending as it's a tradeoff between my working day and conference timings. So, an application that can help me get better reviews of the conference would save me from a tiresome day go wasted due to a bad decision. We do not require our users to end up wasting their valuable time in attending the events that do not meet their target. This gives rise to the need for a tool that would help users utilize their time artfully and judiciously. This application gathers user data and gives an analysis report on same while rating the conference. After analyzing all the essentials, we designed a prototype (figure1) for the application. The most vital components of the UI were added to this prototype. The foremost task was to come up with an appropriate name for the application. Next, a search bar was added to help the user look up for conferences. The ultimate goal of this

application is to display the ratings of conference based on the user feedback on twitter; thus, a "Rating" button was added to the prototype.

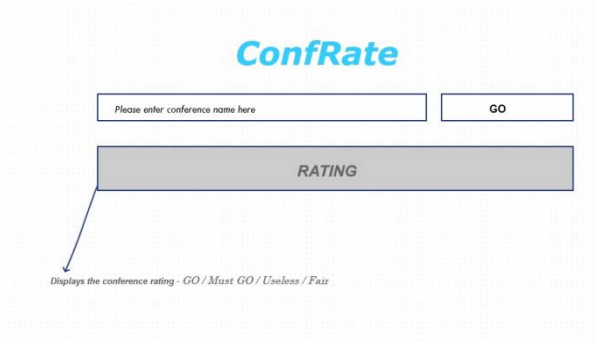


Figure 1 : Application Prototype

III. Implementation Overview

The project work as carried out in the form of sessions wherein each session, we worked on different aspects of the application such as requirement gathering, re-designing the user-interface, team management etc.

a) Session Summary

1. *Session 1 &Session 2*

These sessions were unofficial as during these sessions the team members got to know each other and gathered some background information on the technical aspects of the application to be developed. The sessions that followed, mainly comprised of development and implementation.

2. *Session 3: Requirement Gathering*

In order to identify an appropriate language for the development of Conf Rate, the team did a lot of research on all potential technologies that could be applied. The major task was to come up with a language that is widely used and recognized. Python dominated the list of all possible alternatives. Python's open source license makes it free to use; so anyone can install it. Twython is twitter's Python API that connects it on Oauth (open authentication). Oauth generates a security token which connects it to twitter. Django 1.3.1 has been used as the framework to run python because of its ability to simplify and speed up the fabrication of compound applications. It provides a lightweight web server with MVC framework which could easily integrate with our python script.

3. *Session 4: Planning User-Interface*

We decided to re-design the UI (figure 2) to make it look more appropriate and authentic.

Figure 2 : Re-designed User Interface

Besides having a search button, the new UI has a well-designed homepage. This homepage has added features such as Information about the Standards, Principles, Experience and Reviews. Most importantly, it provides information about the upcoming conferences. This helps the user to get updates about almost all the top-notch conferences.

Further, this UI has been re-designed with the help of the following tools:

i. *HTML 5*

The front end of our application is implemented in HTML5 as it is the latest version and has more features than the previous one. We have used new header and footer tags available in HTML5. Also we have incorporated the favicon using link tag. Html5 has added better functionality to our web pages.

ii. *Javascript*

JavaScript is used in our project to provide basic client side validation. Since the processing is done on the client's browser. It results in less load on the server. Javascript and Jquery (a Javascript library) has been used extensively in our project as it provides a very good support for representation at the front end.

iii. *CSS*

CSS is used throughout the presentation layer along with HTML5 to provide styling support and to make the web pages more presentable. We have incorporated various visual display effects using Css. For example, on the on hover event of the home button, the background color of the button changes because we have changed the Css styling properties on occurrence of that event.

iv. *Jquery*

Jquery has been added as an added functionality to Html5 pages to make our web pages more responsive. Along with that, it has added application scrolling facility to our application.

v. *Highcharts*

Highcharts is a library used to display dynamic and user friendly charts in HTML pages. Highcharts supports HTML5 and is open source. It is a jquery based library hence requires javascipt to be enabled in the user's browser. It has options to display various different chart types like area charts, bar and column charts, pie charts and scatter and bubble charts. We have used pie and column charts (figure 3) in our project. The pie chart is used to display the percentage of positive and negative tweets and the column chart is used to display the sentiment of the last five tweets.

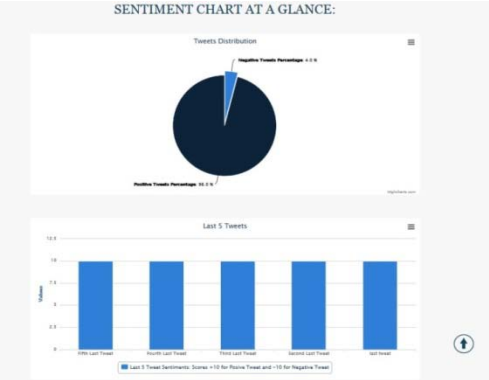


Figure 3 : Pie Chart and Column Chart

vi. *Datatables*

Data Tables are a flexible jquery based javascipt library. They require javascript to be enabled in the client's browser. They also provides various features like sorting data, filtering data on the go, hidden columns, flexible data width and state saving. Besides being trivial to implement, these give a very good look & feel to the front end of an application. Various styling options like adding custom classes, Themeroller support and Theme Forest themes support which are pre-built HTML and CSS templates are supported by Data Tables. We have used Data Tables (figure 4) to display the tweets we are fetching from the twitter api so the end user can also view the tweets.

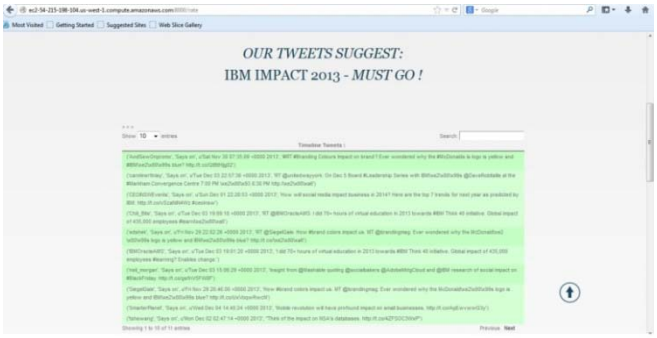


Figure 4 : Data Table displaying the Tweets

4. *Session 5: Execution*

The web framework was implemented using Django. It's written in python and provides MVC architecture. We built some responsive templates in HTML5, CSS etc in our view component so that the user could enter a conference name. The model part of the web server remains empty as there's no database required for our implementation.

In order to connect to twitter to search for the conference name entered by the user, Twython has been used. We used naive Bayes classifier based on Bayes' theorem to classify the tweets as positive or negative at the run time. This classifier can be easily implemented with the use of Python external library and Natural Language toolkit.

The algorithm works in the following manner. The classifier needs to be trained prior to analyzing the sentence. We created two lists of manually classified tweets; positive and negative and tagged each of the sentence with the sentiment it carries. The next step was to extract each of the word in the real time tweets except the stop words that don't carry any sentiment. Then, we defined feature extractor that compares our dictionary of manually classified words (in order of its frequency) with the real time tweet words. Finally, we trained our classifier on the training set in order to create a ready classifier. It detects the sentiment on the basis of the association of positive/ negative words learnt on the training set. We also used data table and High chart to build dynamic interactive data representation of our sentiment.

5. *Session 6: Deployment*

Project has been deployed on Amazon EC2 (Elastic Compute Cloud). Elastic Compute Cloud is a web service offered by Amazon at a minimal cost. We created an instance for our project and established its configuration settings including inbound security for assigning a particular port number for instance to run. To SSH the Amazon web server, we used WINSCP and ran the putty. A new session was opened with instance security key and got launched as a ubuntu user. Thereafter, python, twython, django and nltk were installed on our instance through couple of commands.

We have used screen command to nohup our server and have it run forever in background. We have deployed our website on port number 8000 and following is the url for our website which is already deployed on Amazon EC2 web service:
http://ec2-54-215-198-104.us-west-1.compute.amazonaws.com:8000/

After deployment of the project on the cloud, the images depicting the working of this application are as follows:

Figure 5 shows that the user looks up for conferences and the results related to that particular conference are displayed dynamically.

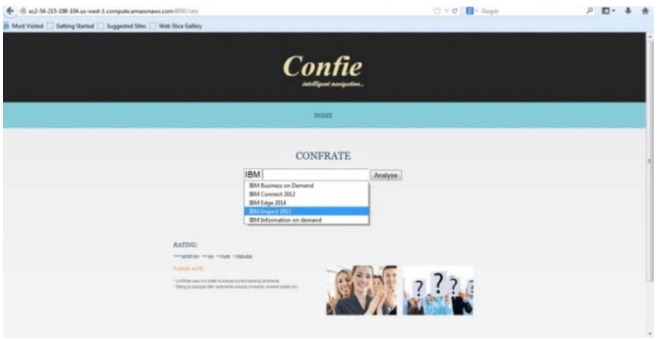


Figure 5 : Conference Search

Next, the user will hit the "Analyze" button and the application will fetch all the tweets related to that conference. These tweets undergo sentiment analysis and a result (figure 6) is displayed by the application. This result is generated in terms of: Must Go, Go, Fair and Useless where Must Go and Go convey a positive feedback, Fair indicates a neutral response and Useless implies that the conference majorly has negative feedback.

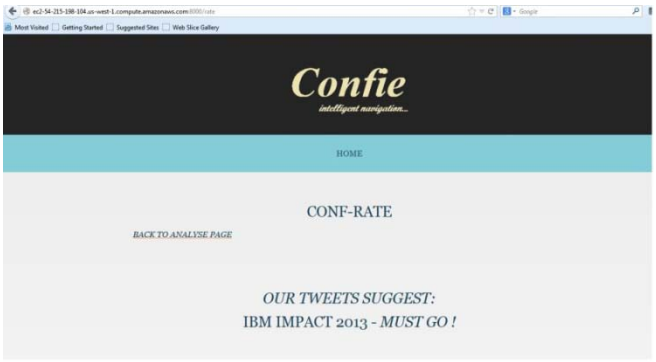


Figure 6 : Rating of the Conference

6. *Session 7: Report*

The last and one of the most crucial part of the project development was the project manual. The manual follows the IEEE format and describes all the essential aspects of the project in a simplified and easy to understand manner. This report covers all the technical and non-technical facets of the tool.

IV. Challenges

Building an application of this potential is not an easy task. Many challenges were faced during this. Among them, the prime challenges were:

The hardest task of all was to develop and implement an algorithm that supports python and is capable of carrying out sentiment analysis.

This issue was solved by using python's natural language toolkit whose functions support the Bayes classifier.

Another challenge was to run the script on a local web server in order to be responsive from the client front end. This task was simplified by the use of Django which is easy to use framework.

Also, it is a tedious task to enable execution of interactive HTML5, CSS and images on the view component of Django. This problem was solved by defining a static path on the server side.

After successfully executing the tool, its deployment was another challenge we faced. We deployed the tool on Amazon web services. The big task was to ensure that the process runs forever on the ec2 instance. The best possible solution to this problem was to make the server run in daemon mode using a screen that would enable a virtual terminal for the processes to pull off in the background.

V. Future Implementations

One of the major challenges of this tool is to fetch the upcoming conferences dynamically. Automatic updating of the conferences that are to be held is a paramount challenge which makes use of cron jobs.

Secondly, data can be gathered and mined in a more optimized manner. We may gather data from yelp for example and use it further to generate ratings. Basically our idea is to integrate website or engines with huge data sets and provide most customized reviews. Opinions from different websites can help give better conclusions about the conferences. More reviews can help analyze sentiments in a better way.

VI. Conclusions

We have named our website as Confie- intelligent navigation. As the name suggests, we are helping our users making an intelligent navigation further through an optimized search for conferences.

We built Conf Rate, which is a python based application that analyses the feedback given by the people who attend conferences. In short, following describe the tool in the best possible manner.

The feedback given by various people who attend conferences is gathered from twitter in form of "tweets". These tweets are analyzed using an algorithm and then they are categorized as "positive" or "negative".

All the tweets related to the search query is fetched at that point of time. Users can have a look at all the tweets that are being tweeted by the people.

Next, the user can view a bar chart that shows the last five recent tweets whether, negative or positive. The pie chart gives an overall percentage of all positive and negative tweets gathered. The user gets the feedback of the conference in term of Must Go, Go, Fair and Useless.

References Références Referencias

1. http://en.wikipedia. org/wiki/Python_(programming_language)
2. http://www.python.org/
3. http://aws.amazon.com/application-hosting/
4. http://en.wikipedia.org/wiki/HTML5#Differences_from_HTML.C2.A04.01_and_XHTML.C2.A01.x
5. http : / /en. wikipedia.org/ wiki/ Django_(web_framework)
6. http://www.w3schools.com
7. http://www.highcharts.com/
8. http://datatables.net/

This page is intentionally left blank

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 14 Issue 1 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Analysis of OLSR, DSR, DYMO Routing Protocols in Mobile Ad-Hoc Networks using Omnet++ Simulation

By Trapti Jain & Savita Shiwani
Suresh Gyan Vihar University, India

Abstract- In present scenario, choosing the routing protocol is vital task in mobile ad-hoc networks. These type of networks is collection of nodes which are connected dynamically and situated without using any infrastructure. There are various types of routing protocols have been implemented such as OLSR, DSR, DYMO, AODV, DSDV, BATMAN etc. These are implemented in specific simulation environments. In this research, an analysis has been done to choose the appropriate routing protocol. A comparison based on relative results is prepared for DYMO, OLSR and DSR protocol. A sample network is simulated to try these three routing protocols over a set of parameters. DYMO and DSR protocols found more difficult and OLSR protocol has better performance in comparison of both DYMO and DSR. This simulation has been carried out using OMNeT++ simulation framework.

Keywords: *AD-HOC network, DYMO, DSR, OLSR, OMNET++, INET.*

GJCST-E Classification : *C.2.2*

Strictly as per the compliance and regulations of:

Analysis of OLSR, DSR, DYMO Routing Protocols in Mobile Ad-Hoc Networks using Omnet++ Simulation

Trapti Jain [α] & Savita Shiwani [σ]

Abstract- In present scenario, choosing the routing protocol is vital task in mobile ad-hoc networks. These type of networks is collection of nodes which are connected dynamically and situated without using any infrastructure. There are various types of routing protocols have been implemented such as OLSR, DSR, DYMO, AODV, DSDV, BATMAN etc. These are implemented in specific simulation environments. In this research, an analysis has been done to choose the appropriate routing protocol. A comparison based on relative results is prepared for DYMO, OLSR and DSR protocol. A sample network is simulated to try these three routing protocols over a set of parameters. DYMO and DSR protocols found more difficult and OLSR protocol has better performance in comparison of both DYMO and DSR. This simulation has been carried out using OMNeT++ simulation framework.

Keywords: AD-HOC network, DYMO, DSR, OLSR, OMNET++, INET.

I. Introduction

Since 1970s, wireless networks have been liked by computer industry. Wireless networks are collection of nodes which provides communication without using any physical link. Mobile wireless networks have two types of variations i.e.

a) Infrastructure networks
b) Infrastructure-less networks

Infrastructure networks are known as collection of fixed nodes using wired gateways. In these networks a mobile node can communicate with nearest base station (which is called as bridge). Mobile node can also move in its geographical limit. If mobile node goes out of its range then it can reconnect using another base station, this process is called handoff.

Infrastructure-less networks are known as collection of nodes which are connected dynamically without any wired link. These types of networks communicate without any rules and laws. These networks have no fixed routers but nodes (work as router) themselves decide the route to transfer and maintain formation from one node to another node. These are also known as mobile Ad-hoc networks.

This research paper goes on to analyze the performance of routing protocols used in mobile ad-hoc networks. Section 2 describes various routing protocols. Section 3 describes OMNet++ simulation framework as well as its features. Section 4 describes network setup and performance analysis. Finally paper ends with significant conclusion.

II. Ad-Hoc Routing Protocols

According to topology information organization Ad-hoc routing protocols are categorized in two different ways:

a) Table-driven routing protocols (proactive)

This type of routing protocol maintains routing information from each node to every other node. Every node maintains one or more routing table to store routing information. When network topology changes then routing information has to be updated at every node. There are various types of protocols which fall in this category such as DSDV, WRP, OLSR, CGSR etc.

b) On-demand routing protocols (reactive)

This type of routing protocol creates routes when desired by source node that's why these are also known as source- initiated routing protocols. When a node wants to transfer information to another then a route discovery process is initiated. After route establishment a route maintenance procedure is called which maintain the route for particular node until route is no longer desired. Examples are AODV, ABR, SSR, TORA, DYMO etc.

Optimal Link State routing protocol

It is abbreviated as OLSR protocol and also known as proactive protocol. It is based on pure link state algorithm. It provides periodic exchange of information to maintain topology changes information at every node. OLSR works in purely distributed manner and suitable for large and dense networks. OLSR performs hop by hop routing technique which means each node uses its most recent information to route a packet in network. It also uses multipoint relaying technique to reduce retransmission of control messages. OLSR supports node mobility that can be traced through its own control messages depended upon frequency of these messages.

Authors α σ: Suresh Gyan Vihar University Jaipur, Rajasthan, India. e-mails: Trapti.it07059@gmail.com, savitashiwani@gmail.com

Dynamic MANET on-demand routing protocol

It is reactive on demand (DYMO) protocol. It offers adaptation to change network topology and determines unicast routes between nodes within network. DYMO stores routes information for each intermediate hop. In DYMO, when a node wants to transmit information to another node then it transmits route request message (RREQ) to all other nodes in range. If an intermediate node receives this message then it appends information about itself and sends messages to all nearby nodes. Receiving node returns route reply (RREP) message unicast to sending node with appending information about itself. DYMO is also designed with future enhancements and uses generic MANET packet and message format and offers ways of dealing with unsupported elements in a sensible way.

Dynamic source routing protocol

It is on-demand routing protocol which is source initiated. DSR protocol has two phases, route discovery and route maintenance. Every node maintains a route cache which stores route information. When a node wants to transmit a packet to another node then firstly it checks in route cache for available path to send packet. If route is still available then message is transmitted otherwise route request packet is broadcasted in route discovery phase. A route reply is generated when the route request reaches the destination node or an intermediate node which contains in its route cache an available route to the destination node. Route maintenance phase is used when an error in route packets occurs. Route error packets are sent by node. When this packet is received by other node then error hop is removed from route cache.

III. Omnet++ Simulation Framework

Sample network has been simulated using OMNet++ 4.2.2. which is available freely for academic use. Scenarios in OMNeT++ are represented by a hierarchy of reusable modules written in C++. Modules relationships and communication links are stored as Network Description (NED) files and can be modeled graphically. Simulations are either run interactively in a graphical environment or are executed as command-line applications.

The INET Framework provides a set of OMNeT++ modules that represent various layers of the Internet protocol suite, e.g. the TCP, UDP, IPv4, and ARP protocols. It also provides modules that allow the modeling of spatial relations of mobile nodes and IEEE 802.11 transmissions between them.

IV. Network Setup

This simulation has been done in windows7. The overall simulation is based on networking simulation framework OMNeT++ (version 4.2.2). In this simulation IEEE 802.11g specifications are preferred. Simulation is run in command-line environment using cmdenv. Nodes are spread randomly over the network without using any mobility model. Message length has been used as only 512 bytes. The playground configuration is used: 1000 m X 800 m with 10, 20 and 30 nodes. The simulation time is decided as 100 seconds. Packets have been transmitted randomly with uniformly distributed speed (0 to 25 seconds). When packet is reached to its destination node then another packet is ready to transmit to destination node which is again randomly chosen.

Fig 1 shows network design in simulation progress of network in tcl/tkenv graphical environment with 30 nodes.

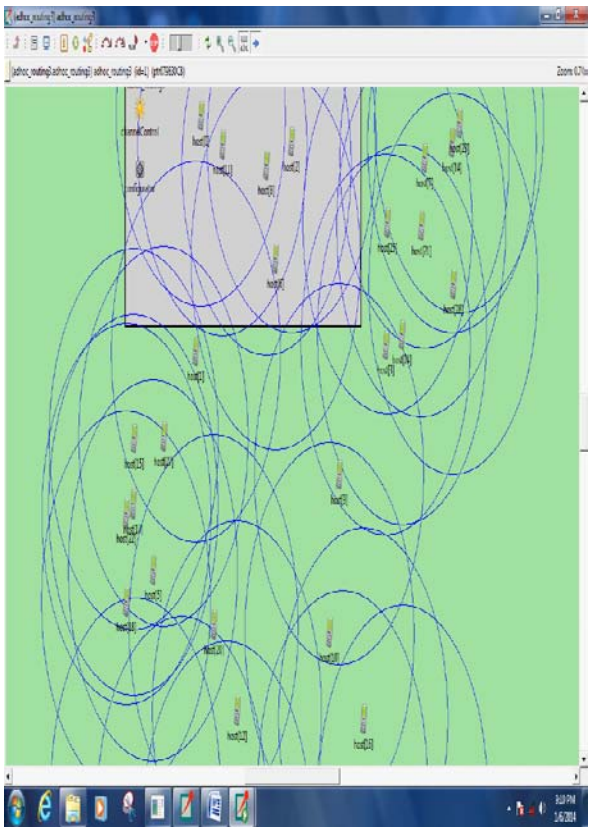


Figure 1 : network design with nodes =30

a) Performance analysis

There are following parameters that have been analyzed as explained below:

Throughput - Throughput is a parameter, which is measured in either bits/sec or data packets per sec. Throughput is rate of successfully delivered packets over the networks. In ad-hoc networks data packets are delivered using nodes.

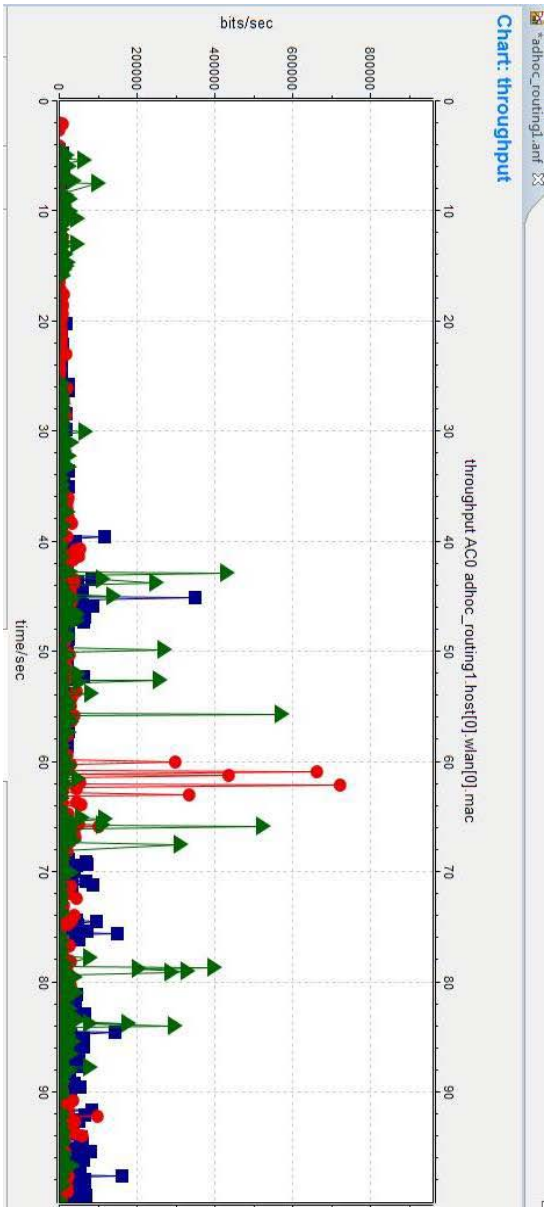


Figure 2 : (a) throughput at host #0 when nodes=10

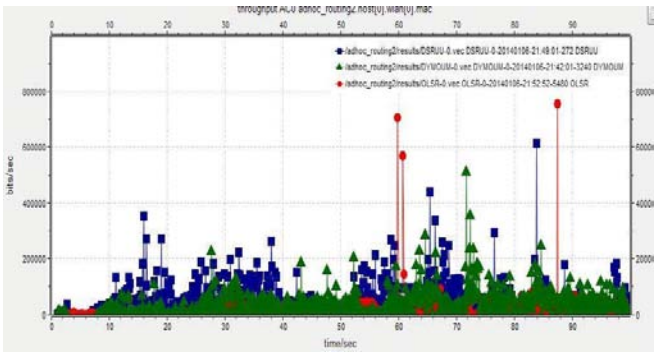


Figure 2 : (b) throughput at host #0 when nodes =20

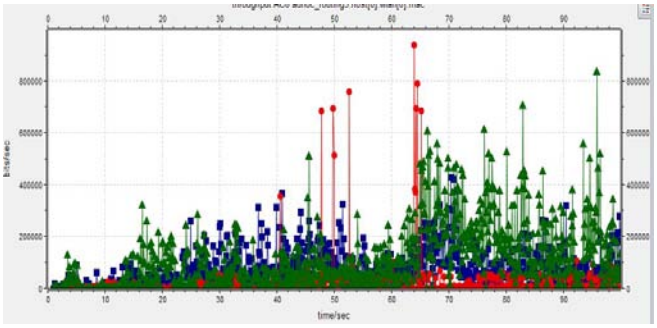


Figure 2 : (c) throughput at host #0 when nodes =30

Figure 2 (a,b,c) shows the graphs generated from the vector data for three protocols during simulation time period. There is comparative analysis of three protocols but conclusion cannot be made clear according to this parameter.

Jitter

Jitter is a parameter which is known as variation in latency. It is measured as variability over the time of packet latency across the network. Its standard term is packet delay variation (PDV). In this simulation it has been measured as mean value for three protocols using window size=10. When packet delay variation is zero then value of jitter will be zero.

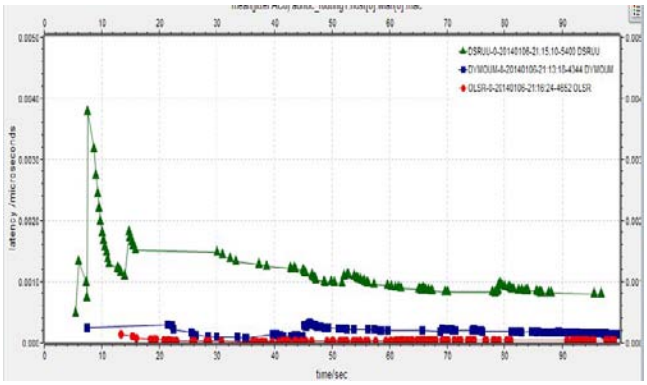


Figure 3 : (a) jitter at host #0 when nodes=10

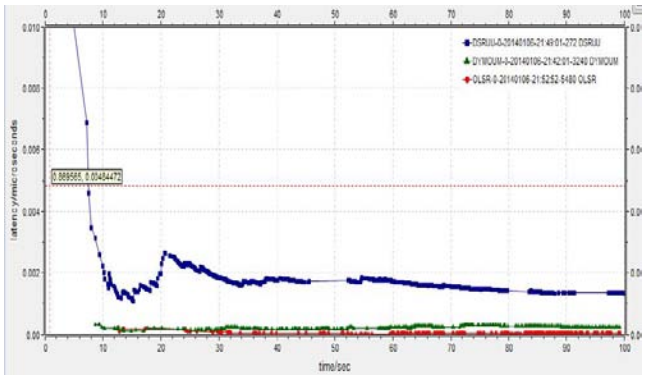


Figure 3 : (b) jitter at host #0 when nodes=20

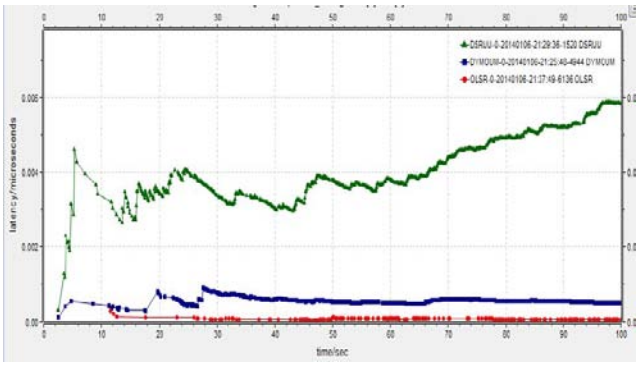


Figure 3 : (c) jitter at host #0 when nodes=30

Figure 3 (a,b,c) shows the graph generated from vector data during simulation. Initially it is shown that when no of nodes 10 then DSR protocol has value of jitter larger then OLSR and DYMO protocol. OLSR has minimum value of jitter approximately zero in comparison of others regardless no of nodes. OLSR is better in this case.

Now there has been made a comparison table for various parameters as shown below:

NO of nodes	Parameters	OLSR	DSR	DYMO
10	End to end delay	Low	high	Very high
	Throughput	Very high	high	high
	Jitter	Very low	high	low
	No of collisions	Low	high	high
	SNIR	Low	high	low
	Dropped packets by queue	Zero	high	zero

NO of nodes	Parameters	OLSR	DSR	DYMO
30	End to end delay	low	high	Very high
	Throughput	Very high	high	low
	Jitter	Very low	high	low
	No of collisions	low	Very high	high
	SNIR	Very low	Very high	high
	Dropped packets by queue	zero	high	zero

V. Conclusion

In this research paper, three protocols are evaluated such as OLSR, DYMO, DSR. The comparison has been made over a set of parameters increasing the no of nodes during each simulation. In conclusion, it has been shown that OLSR is better than DYMO and DSR routing protocol.

References Références Referencias

1. Nitin Kumar, Dr. Parvez Ahmed Alvi, Dr. Ajay Singh yadav, Ms. Anupam Swami "A Study of Routing Protocols for Ad-hoc Network".
2. Samyak Shah, Amit Khandre, Mahesh Shirole and Girish Bhole "Performance Evaluation of Ad Hoc Routing Protocols Using NS2 Simulation".
3. Isabel Dietrich, Christoph Sommer and Falko Dressler "Simulating DYMO in OMNeT++".

Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 14 Issue 1 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Modified TCP for Time Critical Applications

By Abhay Kumar
JB Institue of Engineering and Technology, India

Abstract- A network is defined to be a congested network if the load on the network exceeds the capacity of the network. The traditional congestion control technique of slow-start and AIMD was adopted when the aim was more on the stability of the Internet. But as more and more time critical applications such as multimedia applications are being used, we need alternate technique that reduces the drastic fluctuations of window size present in the existing technique. Thispaper proposes a techniquefor fast delivery of packet for a time critical application. It reduces the packet overhead and time compared to existing slow-start and AIMD technique. The proposed technique uses information or intelligence from the unexpected packet received. It is a fine modification of the existing slow-start and AIMD technique by adapting them for time critical applications. We propose modification at both thesender and the receiver hosts without modifying anything in the intermediate hosts of the network. Extensivesimulation show that proposed technique reduces congestion in the network by reducing both packet overhead and time compared to the traditional slow-start and AIMD technique and delivers the packets in timely manner than the existing techniques.

Keywords: *network protocols, TCP, congestion control, slow-start, aimd.*

GJCST-E Classification : *C.2.5*

Strictly as per the compliance and regulations of:

Modified TCP for Time Critical Applications

Abhay Kumar

Abstract- A network is defined to be a congested network if the load on the network exceeds the capacity of the network. The traditional congestion control technique of slow-start and AIMD was adopted when the aim was more on the stability of the Internet. But as more and more time critical applications such as multimedia applications are being used, we need alternate technique that reduces the drastic fluctuations of window size present in the existing technique. Thispaper proposes a techniquefor fast delivery of packet for a time critical application. It reduces the packet overhead and time compared to existing slow-start and AIMD technique. The proposed technique uses information or intelligence from the unexpected packet received. It is a fine modification of the existing slow-start and AIMD technique by adapting them for time critical applications. We propose modification at both thesender and the receiver hosts without modifying anything in the intermediate hosts of the network. Extensivesimulation show that proposed technique reduces congestion in the network by reducing both packet overhead and time compared to the traditional slow-start and AIMD technique and delivers the packets in timely manner than the existing techniques.

Keywords: network protocols, TCP, congestion control, slow-start, aimd.

I. Introduction

The Internet is a global network of interconnected computers which allows individuals and organizations around the world to communicate and share information with each other. This demand has natural fluctuation; therefore, the Internet performance is largely governed by it, leading to possible congestion which occurs when resource demands exceed the capacity of the network. Due to the explosive growth of the Internet and increasing demand for multimedia applications like voice over IP, real-time video streaming, IPTV and financial transactions, the issue of congestion has received tremendous attention from academia and industry. Transmission of real-time multimedia applications typically has large bandwidth, small delay and low-loss requirements. However, the current Internet does not guarantee any quality of service (QoS) as it is based on best-effort service model of IP [1]. A network is said to be congested from the perspective of a user if the service quality noticed by the user decreases because of an increase in network load. The goal of congestion control mechanisms is simply to use the network as efficiently as possible, that is, attain the highest possible throughput while maintaining a low loss ratio and small delay. Congestion must be avoided because it leads to queue growth and queue growth leads to delay and loss [2

As the network grew, it was clear that unrestricted data transfer by many users over a shared resource, i.e., the Internet, could be bad for the end users; excess load on the links leads to packet loss and decreases the effective throughput. This kind of loss was experienced at a significant level in the '80s and was termed congestion collapse [5]. Thus, there was a need for a protocol to control the congestion in the network, i.e., control the overloading of the network resources. It led to the development of a congestion control algorithm for the Internet by Van Jacobson [5]. This congestion control algorithm was implemented within the protocol used by the end hosts for data transfer called the Transmission Control Protocol (TCP).

There are several different flavors of TCP congestion control, each of which operates somewhat differently. But most of the versions of TCP are window-based protocols, wherein the idea is that each user maintains a number called a window size, which is the number of unacknowledged packets that are allowed to be sent into the network. Any packet from the new window can be sent only when an acknowledgment for the last packet in the previous sent window is received by the sender. TCP adapts the window size in response to congestion information. The window size is increased if the sender determines that there is excess capacity present in the route and decreases if the sender determines that the current number of in-flight packets exceeds the capacity of the route. The exact means of determining whether to increase or decrease the window size is what determines the difference between the congestion control mechanisms of different TCP flavors. The most commonly used TCP flavors used for congestion control in the Internet today are Reno and New Reno [12]. Both of them are updates of the TCP-Tahoe, which was introduced in 1988[5]. Although, they vary significantly in many regards, the basic approach to congestion control is similar. The idea is to use successful reception of packets as an indication of available capacity and dropped packets as an indication of congestion. In most cases, eachtime the destination receives a packet, it sends an acknowledgement (also called ACK) asking for the next packet in sequence to be transmitted. When an acknowledgment for a windowis received, the protocol increases its window size. However, on reception of three duplicate acknowledgments or dupacks (i.e., four successive

Author: JB Institue of Engineering and Technology, Hyderabad, India. e-mail: abhay1880@gmail.com

identical acknowledgments) by the sender is taken by it as an indication that packet has been lost due to congestion. In case the source does not receive any acknowledgement for a finite time (RTO [13]), it assumes that all unacknowledged packets have been lost. In both the cases the source aggressively proceeds to cut down the window size and retransmit the lost packets.

TCP Vegas improves upon TCP Reno through three main techniques. The first is a new retransmission mechanism where timeout is checked on receiving the first duplicate acknowledgment, rather than waiting for the third duplicate acknowledgment, and results in a more timely detection of loss. The second technique is a more prudent way to grow the window size during the initial use of slow-start when a connection starts up, and it results in fewer losses. The third technique is a new congestion avoidance mechanism that corrects the oscillatory behavior of Reno. The idea is to have a source estimate the number of its own packets buffered in the path and try to keep this number between α (typically 1) and β (typically 3) by adjusting its window size. The window size is increased or decreased linearly in the next round-trip time according to whether the current estimate is less than α or greater than β. Otherwise the window size is unchanged. The rationale behind this is to maintain a small number of packets in the pipe to take advantage of extra capacity when it becomes available. A source periodically measures the round-trip queuing delay and sets its rate to be proportional to the ratio of its round-trip propagation delay to queuing delay, the proportionality constant being between α and β. Hence, the more congested its path, the higher the queuing delay and the lower the rate. The Vegas source obtains queuing delay by monitoring its round-trip time (the time between sending a packet and receiving its acknowledgment) and subtracting from it the round-trip propagation delay [7].

In today's Internet, real-time applications such as VoIP, videoconferencing and on-line gaming mostly use RTP over UDP or UDP alone to transport data. Because these protocols are unresponsive to congestion events, the growing popularity of applications that use them endangers the stability of the Internet. So, to make it possible that real-time applications are widely adopted, common congestion control mechanisms suitable for real time multimedia are expected to be deployed[3] [4].

The existing techniques does not use any information or intelligence from the unexpected packet received, unexpected packets are simply discarded. The proposed techniques tries to retrieve information based on the unexpected packet received and perform the congestion control accordingly.

The remaining paper is organized as follows: Section II explains the system or network model used in this paper. Section III describes our proposed Unexpected Packet based Congestion Control (UPCC) Technique. Section IV presents the simulation results that demonstrate our proposed UPCC technique reduces congestion in the network compared to traditional slow-start and AIMD technique. Finally Section V concludes the paper.

II. System Model

This paper considers a realistic computer network consisting of several sources and destinations connected via multiple routers and links. The source (sender) communicates to the destination (receiver) in form of packets. The series of routers and links that a packet follows from the source to destination is called a route. A pair of sender and receiver may be connected via multiple routes. This network is represented in the Figure 1.

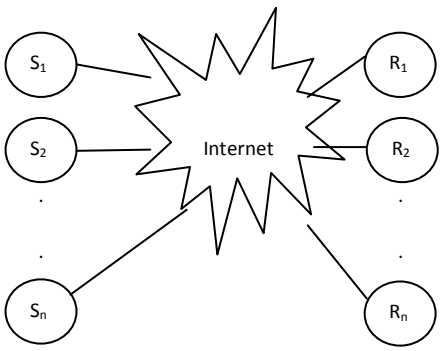


Figure 1 : Network model

For simplicity of the explanation, we consider only a pair of sender (S) and receiver (R) connected via multiple routes, as shown in Figure 2. The sender and the receiver may be running multiple different applications. However, the packets of the application are transmitted using the first come first serve policy. The connection is established using three-way handshake as in case of existing TCP. However, this paper proposes few modifications in this phase also to make the subsequent transmissions congestion aware.

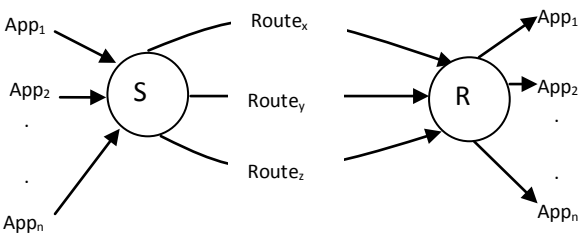


Figure 2 : Simplified network model

TCP operates in two distinct phases. When file transfer begins, the window size is 1, but the source rapidly increases its transmission window size so as to reach the available capacity quickly. Let us denote the window size by W. The algorithm increases the window size by 1 each time an acknowledgement for a packet indicating success is received. This is called the slow-start phase. Since one would receive acknowledgements corresponding to one window's worth of packets in an RTT [13], and we increase the window size by one for each successful packet transmission, this also means that (if all transmissions are successful) the window would double in each RTT, so wehave an exponential increase in rate as time proceeds. Slow-start refers to the fact that the window size is still small in this phase, but the rate at which the window is increased is quite rapid. When the window size either hits a threshold, called the slow-start threshold (ssthresh) or the transmission suffers a loss (immediately leading to a halving of window size), the algorithm shifts to a more conservative approach called the congestion avoidance phase. When in the congestion-avoidance phase, the algorithm increases the window size by 1 every time feedback of a successful packet transmission in the corresponding window is received. When a packet loss is detected by the receipt of three dupacks, the slow-start threshold (ssthresh) is set to half of the current window i. e TCP Reno cuts its window size by half (W ← W/2) and algorithm enters additive increase phase where it start sending segments from current window onwards. Thus, in each RTT, the window increases by one packet i.e., a linear increase in rate. Protocols of this sort where increment is by a constant amount, but the decrement is by a multiplicative factor are called additive increase multiplicative decrease (AIMD) protocols. When packet loss is detected by a timeout, the slow-start threshold (ssthresh) is set to half of the current window and the algorithm enters the slow-start phase i.e., it start sending from 1 packet onwards. Let us call the congestion window at time t as W (t). This means that the number of packets in-flight is W (t). The time taken by each of these packets to reach the destination, and for the corresponding acknowledgement to be received is RTT. The RTT is a combination of propagation delay and queuing delay. A window-based congestion control scheme defines one control rule for window increase, and another rule forwindow decrease. AIMD uses the following control rule [19]:

$$\text{Increase: } W_{t+1} \leftarrow W_t + \alpha, \quad \alpha > 0$$

$$\text{Decrease: } W_t \leftarrow W_t - \beta W_t, \quad 0 < \beta > 1$$

Where α and β refer to the additive increase constant and multiplicative decrease constant β respectively. The standard TCPuses the value of these constants α and β as 1 and 0.5 respectively.

This subsection provides the definition of several terms and the notations that will beused throughout the remainder of this paper.

- *SYN:* To establish a connection, TCP uses a three-way handshake. Synchronize (SYN) [9] packet is the first control packet sent for the three-way handshake by the sender wishing to establish the TCP connection.
- *ACK:* An acknowledgement (ACK) [14] is a control packet used between communicating processes or computers to signify receipt of receiving a data packet, and it is a part of a communication protocol. For example, ACK packets are used in the Transmission Control Protocol (TCP) to acknowledge the receipt of SYN packets while establishing a connection in three-way handshake, and acknowledge the receipt of data packets while a connection is in data transfer phase.
- *SS-AIMD:* In the Slow-Start (SS) [5] [8] and Additive Increase Multiplicative Decrease (AIMD)[5] [14] algorithm, when a TCP connection first starts, the slow-start phase initializes a congestion window to one packet and transmits. After receiving acknowledgement from the receiver, the window increases by one packet for each acknowledgement returned. After successful transmission of these two packets and acknowledgements received, the window is increased to four packets and so on, doubling from there up to a threshold known as slow-start threshold (ssthresh). After slow-start threshold, the algorithm enters into additive increase multiplicative decrease (AIMD) phase where window increases by one packet for successful transmission of all the packets in the window i.e., additive increase. In this phase, the transmission rate slows down to avoid congestion. But whenever a packet is lost, the sender immediately sets its transmission window to one half of the current window size i.e., multiplicative decrease.
- *ssthresh:* Slow-start threshold (ssthresh)[2] is a point where slow-start phase ends and additive increase multiplicative decrease (AIMD) phase starts.
- *dupacks:* When receiver receives a TCP packet with a sequence number higher than the expected one (out of turn packet). The receiver sends an immediate ACK with the Acknowledgement field set to the Sequence number the receiver was expecting. This ACK is a duplicate of an ACK (dupacks) [2] which was sent previously. This is done to update the sender with regards to the missing TCP packets.
- *rwnd:* Receiver advertised window (rwnd)[10] or receiver queue capacity is the most recent advertised window that contains the number of packets a receiver can process. This is one of the

two variables which affect how much unacknowledged data a sender can send; the other variable is congestion window. The receiver advertised window is the value of the window field in a TCP packet header.

- *cwnd:* Congestion window (cwnd)[12] is a TCP state variable maintained at the sender that limits the amount of data a TCP can transmit without facing congestion through the network. At any given time, a TCP transmit minimum of congestion window and receiver advertised window.
- *TCP:* The Transmission Control Protocol (TCP)[14] is used as a highlyreliable host-to-host protocol between hosts in packet-switched computercommunication networks, and in interconnected systems of such networks.TCP is a connection-oriented, end-to-end reliable protocol designed to fit into a layered hierarchy of transport layer protocolswhich support multi-network applications. The TCP provides for reliable inter-process communicationbetween pairs of processes in host computers attached to distinct but interconnected computer communication networks.
- *UDP:* The User Datagram Protocol (UDP)[15] is defined as a datagram mode of packet-switched computer communication in the environment of an interconnected set of computer networks. This protocol assumes that the Internet Protocol (IP) [16] is used as the underlying protocol. User Datagram Protocol is unreliable connection-less protocol used at transport layer
- *IP:* The Internet Protocol (IP)[17] is designed for use in interconnected systems ofpacket-switched computer communication networks. The internet protocol provides fortransmitting blocks of data called datagram from sources to destinations. The internet protocol also provides forfragmentation and reassembly of long datagram, if required, fortransmission through "small packet" networks. Internet Protocol is unreliable connection-less protocol used at network layer
- *RTP:* The real-time transport protocol (RTP)[18] provides end-to-end network transport functions suitable forapplications transmitting real-time information, like audio, video ordata, over multicast or unicast network services. RTP does not provide resource reservation and also does not guarantee quality-of-service for real-time services. This transport protocol is also augmented by another real-time control protocol (RTCP) to allow monitoring of the data delivery in amanner scalable to large multicast networks, and to provide minimalcontrol and identification functionality. RTP and RTCP are designedto be independent of the underlying transport and network layers.
- *VoIP:* Voice over Internet Protocol (VoIP) [3] is a mechanism that allows telephone calls to be made over computer networks like the Internet. VoIP converts analog voice signals into digital data packets and supports real-time, two-way transmission of conversations using Internet Protocol.
- *IPTV:* Internet Protocol television (IPTV)[3] is the process of transmitting and broadcasting television programs using the Internet protocol suite over a packet-switched network such as the Internet, instead of being delivered through traditional terrestrial, satellite signal and cable television formats.
- *RTO:* The retransmission timeout (RTO) [13] is aretransmission timer used by the Transmission Control Protocol to ensure data delivery in the absence of anyfeedback from the remote data receiver. The duration of this timeris referred to as RTO. The retransmission timeout timer is used for retransmissions of lost or delayed packet.
- RTT (Δ): Round trip time (RTT)[13] is the length of time it takes for a packet to be sent and the length of time it takes for an acknowledgment of that packet to be received
- *QoS:* Quality of service (QoS) [2] is the ability to provide different priority to different applications, users, or data flows i.e., it guarantees a certain level of performance to a data flow. Quality of service guarantees are important if the network capacity is insufficient, especially for real-time multimedia applications such as voice over IP, online games and IPTV, since these applications often require fixed bit rate and are delay sensitive. A best-effort network like Internet does not support quality of service.

Table 1 : Notations.

X_E	The number of windows and hence number of acknowledgements used in the existing slow-start and AIMD technique
X_P	The number of windows and hence number of acknowledgements used in the proposed Unexpected Packet based Congestion Control technique
T_E	The total time required to transmit an application in the existing slow-start and AIMD technique
T_P	The total time required to transmit an application in the proposed Unexpected Packet based Congestion Control technique
W	Window size
W_s	Window start
W_{end}	Window end
Δ	Round trip time
δ	Time required to transmit consecutive packets in a window

III. Proposed Unexpected Packet Based Congestion Control (UPCC) Technique

The proposed technique is a fine modification of the existing slow-start and AIMD technique by adapting it and making congestion aware. We propose modification at both the sender and receiver hosts without modifying anything in the intermediate hosts of the network. The proposed modification can be described in the form of a dialogue between sender and receiver from initiation to the termination of a connection.

At sender side:

Whenever a sender host wants to communicate it will send a SYN (i) packet to the receiver host expressing its desire to communicate as in existing technique [8] [9]. On sending the SYN(i) packet the sender will start a timer based on RTT within which it should ideally receive an ACK (i+1) packet from the receiver. This can be seen in Figure 3. In case, he does not receive an ACK (i+1) packet, he assumes that there is congestion in the network and therefore it retransmit SYN (i) packet with doubled RTT. This information about congestion is stored in a separate variable 'C' that will be used in data transfer stage, i.e., it set C=1. This communication can be seen in Figure 4.

Algorithm for three-way handshake at sender

start

Send a SYN message and start a timer
Wait for an ACK
If timer expires and no ACK received
C=1
Resend the SYN message with RTT=2RTT

stop

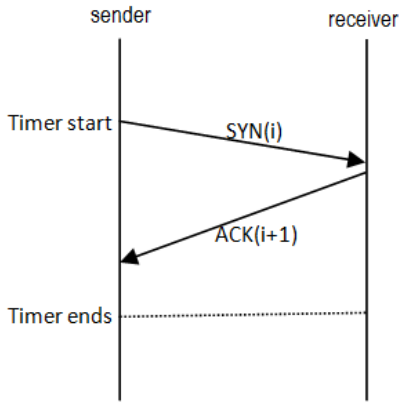


Figure 3 : Three-way handshakein an ideal condition.

At receiver side:

On receiving a SYN(i) packet it will send an ACK(i+1) packet containing its available queue capacity 'rwnd' together with its own SYN(j) and set C=0 to inform its readiness for communication and no congestion perceived so far. To complete the three-way handshake of TCP connection it starts its timer waiting for an ACK (j+1) from sender for his SYN (j) as shown in Figure 5. However, if it receives unexpected duplicate SYN (i) message or no ACK (j+1) within its RTT, it indicates that its ACK (i+1) or ACK (j+1) was lost and hence congestion may be present. It responds to this new SYN (i) received or RTT time out by

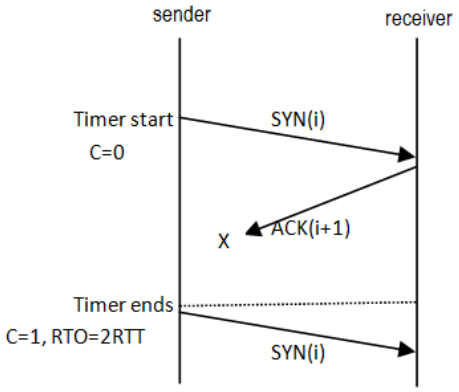


Figure 4 : Three-way handshakewhen ACK from receiver is lost.

retransmitting with a packet containing SYN(j), ACK(i+1), and rwnd. This information about congestion is stored in a separate variable 'C' that will be used in data transfer stage, i.e., it set C=1. This communication can be seen in Figures 6 and 7.

Algorithm for three-way handshake at receiver

start

If (SYN message received)
Send ACK+SYN message and start a timer
If timer expires and no ACK received or duplicate SYN (i) is received
C=1
Resend the ACK+SYN message with RTT=2RTT

stop

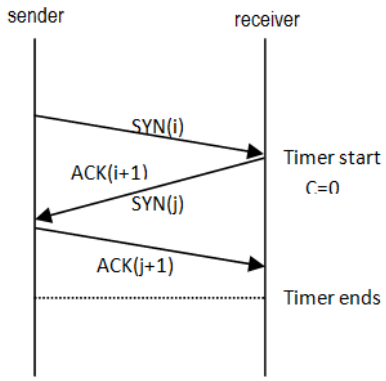


Figure 5 : Three-way handshake in an ideal condition

After the three-way handshake is completed, proposed algorithm enters the data transfer phase. However, during the handshake if no timer expires or no duplicate SYN or ACK packets are received, the

proposed technique presumes network to be congestion free. Thus, it advocates an aggressive start wherein the window size is set to be equalto the receiver queue capacity 'rwnd'. On the other hand, a congestion may be perceived when C=1 at either the sender or receiver side. In such case we follow the same existing

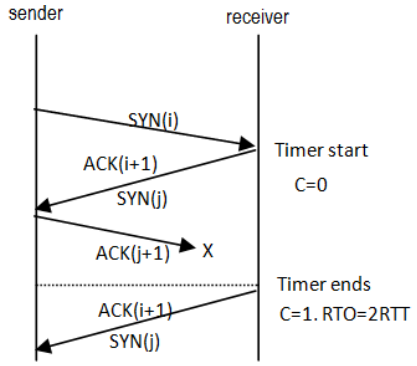


Figure 6 : Three-way handshake when ACK from sender is lost.

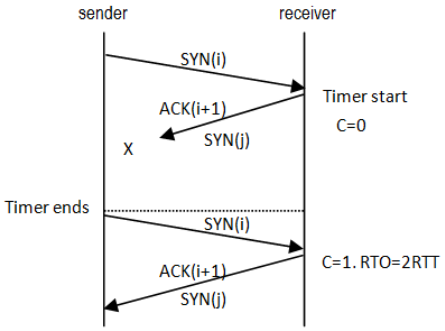


Figure 7 : Three-way handshake when receiver receives duplicate unexpected SYN(i).

slow-start and AIMD technique[5] [8] for selecting the window size. After the selection of window size is made, the data transfer phase is initiated by the sender and the dialogue continues as follows:

Algorithm for window selection

start

If C=0

Window Size = rwnd

// we apply aggressive start i.e., it does not depend on cwnd as per standard TCP [5]

If C=1

Window Size = min (cwnd, rwnd)

// we apply the standard TCP rule i.e., slow-start with AIMD [8]

stop

At sender side:

The sender will start sending the packets up to the window size (W_s, W_{end}) but it doesn't expects any ACK till it completes sending the entire window. In other words, it expects one ACK (w_{end}) per window. In ideal condition it will receive the ACK (w_{end}) and assumes no congestion C=0 and will adjusts the window as per the policy defined above, in the algorithm for window selection.

At receiver side:

On receiving the ACK (j+1) with the window size it will set its window and will wait to receive the data packets. When requisite packets arrive it acknowledges them by sending ACK (w_{end}) for the same. However, at any point of time, if it feels overloaded or underloaded, it will send its updated queue capacity 'rwnd' to the sender piggybacking with ACK(k) where k-1 is the last packet accepted from the sender.

At sender side:

If it receives an unexpected ACK (k) (as it expects only ACK (w_{end}) for any window) then it will simply slides the window such that it starts with the first unacknowledged packet, i.e., packet with sequence number k. Further, it adjusts the window according to the new 'rwnd' suggested by the receiver. Thus, on receiving one unexpected ACK (k) the sender simply slides and adjust the window size and again expects one ACK (w_{end}) within the RTT of the new window. This communication can be seen in Figure 8 where 'k = n+3' and new 'rwnd = 12'.

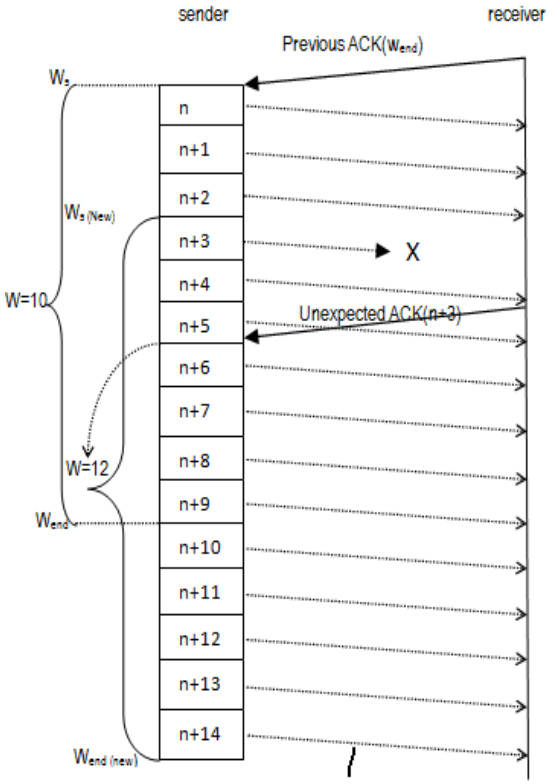


Figure 8 : Data transfer phase when sender receives an unexpected ACK with new rwnd

At receiver side:

The above dialogue presumes that no congestion exists and hence, no packet loss occurs. However, if the receiver finds an out of turn packet it indicates that the intermediate packet/s could be lost. In such case it will send an ACK (k) with current 'rwnd' for the last in order packet i.e k-1 received. It will also slide its window but it does not expect a retransmission of the intermediate packet/s as they may be delayed. However, if it further receives second out of the turn packet it presumes that intermediate packet/s is lost. It sends a duplicate ACK (k) with current 'rwnd' and starts a timer based on RTT within which it should receive the lost packet. In case it does not, it will resend an ACK (k). This communication can be seen in Figure 9.

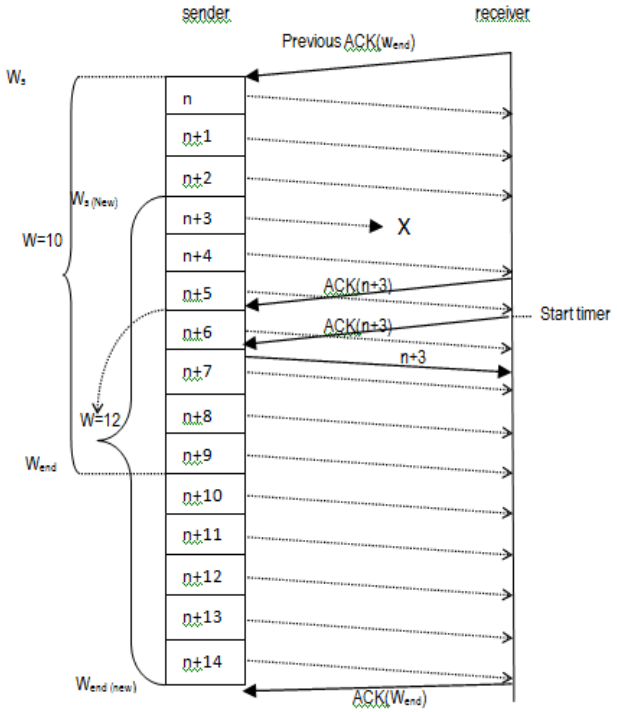


Figure 9 : Data transfer phase when sender receives duplicate unexpected ACK

At sender side:

On receiving the first unexpected ACK (k), the sender simply slides the window as was discussed in Figure 8. But if it receive a duplicate ACK (k), i.e., two ACK (k) it indicates that mild congestion is present in the network. This assumption of mild congestion is based on the understanding between sender and receiver that two duplicate acknowledgements will be send by the receiver only when the receiver receives two out of turn packets. Therefore, it must retransmit only that missing k^{th} packet and continue with sending the packets from first non-transmitted packets in the current window and expect the ACK (w_{end}) for the current entire window. This communication can be seen in Figure 9.

At receiver side:

On receiving the missing packet, it will place it in order and continue receiving till the end of window. If all the packets arrive, the receiver will send the ACK (wend). However, if it misses another packet in the same window, it indicates that the congestion is increasing and it will send the duplicate ACK (j) with 'rwnd'=rwnd/2 as shown in Figure 10.

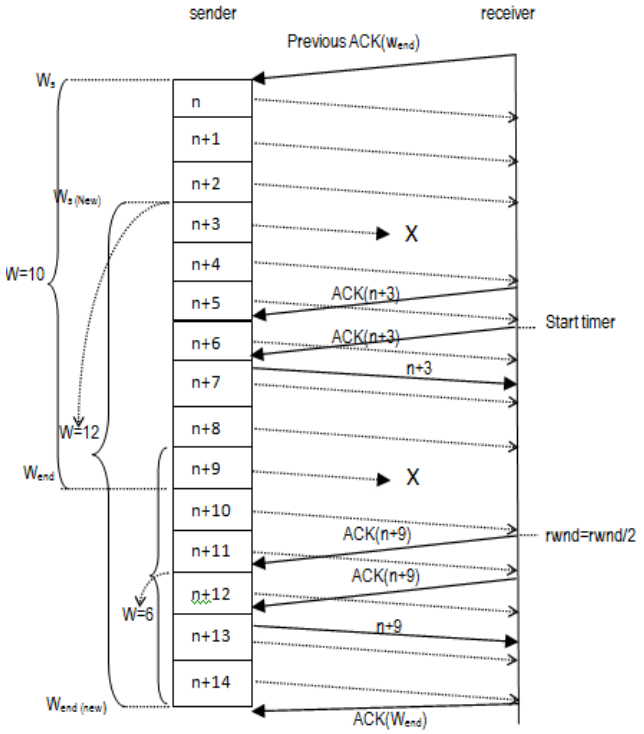


Figure 10 : Data transfer phase when sender receives second pair of unexpected ACK with reduced rwnd.

At sender side:

If sender receives another pair of unexpected ACK (j) in its current window, it indicates that the second packet in the same window has been lost implying that window size is too big. In such scenario the sender will slide the window to the first unacknowledged packet and retransmit the missing packet. It will also reduce its transmission window as indicated by the receiver to half. This communication can be seen in Figure 10.After transmission of the entire window the sender waits for RTT time to receive the acknowledgement ACK (wend). If it receives ACK (wend) within the stipulated time then he assumes that the network is congestion free and continues with the next window. However, if ACK wend) is not received within the RTT the sender presumes high congestion in the network. It retransmits the first packet in the window as shown in Figure 11, and starts the timer with RTT time as perexisting slow-start and AIMD algorithm [5] [8].

At receiver side:

If retransmission of a packet which is not asked by the receiver i.e., unexpected packet is received. The receiver will transmits the ACK (k) where k-1 is the last in order packet received. As demonstrated in the Figure 11, when the sender retransmits the first packet of the last unacknowledged window i.e., Ws=n+9 when it does not receive ACK (n+15) i.e., ACK (W_{end}) within its RTT, the

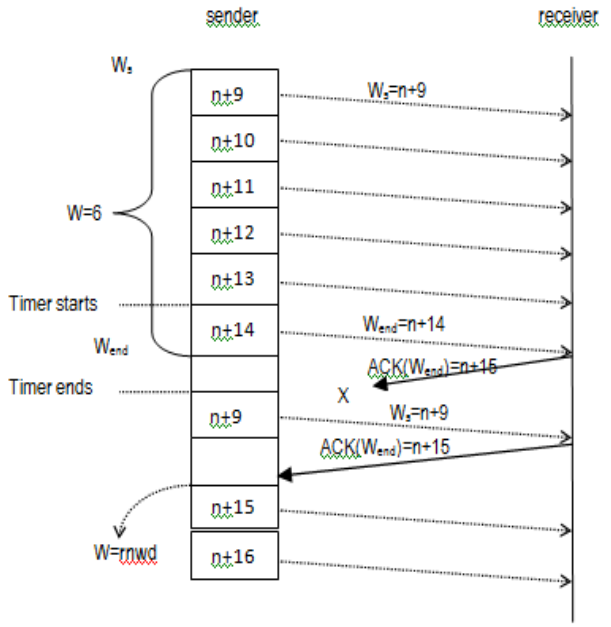


Figure 11 : Data transfer phase when ACK(W_{end}) for complete window is lost

receiver will respond by retransmitting the ACK (n+15) i.e., ACK (W_{end}) indicating the receipt of the complete window n+9 to n+14. By doing this the receiver avoids the retransmission of the remaining packets in the last unacknowledged window i.e., n+10 to n+14. For large window this is substantial reduction in retransmission improving the throughput of the network and reducing congestion.

At sender side

It may receive a unexpected delayed ACK (Wend) in response to Ws retransmitted by it for the previous window. As shown in Figure 11, it receives ACK (n+15) i.e., ACK (W_{end}) in response to its retransmission of packet W_s=n+9. Theexisting techniques [8] will discard this ACK (W_{end}). However, in the proposed technique it is not simply discarded but is used to indicate mild congestion and the previous window packets have been received without any problem. Thus, sender should stop retransmission and recover from slow-start phase by sliding the window up to the first unacknowledged packet and continue with the original window size.

Both sender and receiver utilize the congestion information received for one connection over all other connections made by them leading to recovery from the congestion by the network.

The proposed Unexpected Packet based Congestion Control (UPCC) technique is illustrated with the following motivational example which will illustrate the limitations of existing techniques.

Let an application have 1024 packets and as considered by the authors in [10] [11], the slow-start threshold (ssthresh) as 40 packets and receiver advertised window (rwnd) as 50 packets, we also consider the same. We estimate the packet overhead gain and the time gain for the existing slow-start AIMD technique and the proposed UPCC techniques as follows.

a) Existing SS-AIMD [8] technique in congestion free network

The existing slow-start technique [5] will initially set the window as one packet. When its corresponding ACK arrives, the source sets the corresponding window to two packets. It then transmits two packets. On receiving the two corresponding ACK, it sets the window size to four and so on. Therefore, the slow start technique increases the window size from 1 exponentially up to ssthreshold of 40 packets, forming a geometric progression of 1, 2, 4, 8, 16, and 32. From ssthreshold to rwnd, it will perform additive increase as arithmetic progression of 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 and 50. Beyond rwnd, the current size of the window cannot increase because it has to be minimum of cwnd and rwnd, thus it remains constant at rwnd. The existing technique during the slow start phase will expect an acknowledgement per packet, while in the subsequent phase only one ACK per window will be received. Hence, apart from the 1024 data packet additionally 84 ACK packets will be required.

However, to give the existing technique a fair chance we assume that only one ACK per window is required. Mathematically, the minimum number of ACKs required to transmit the application of 1024 packets in congestion free network is

$$X_E = n_{GP} + n_{AP} + \left\lceil \frac{Number\ of\ packets\ to\ be\ transfered\ -(S_{AP}+S_{GP})}{rwnd} \right\rceil$$

Where

$n_{GP} = \lceil log_2 ssthresh \rceil$

n_{AP} = (rwnd-ssthresh)+1

$S_{AP} = \frac{(rwnd - ssthresh+1)(ssthresh+rwnd)}{2}$

$S_{GP} = 2^{\lceil log_2 ssthresh \rceil}-1$

Thus, the variation in window size will be as 1, 2, 4, 8, 16, 32,40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50,50, 50, 50, 50, 50, 50, 50, 50, 50and 16. In other words, additional X_E=27 number of acknowledgements must be sent by the receiver to acknowledge the correct

receipt of each window, causing the packet overhead in the network.

The time required by the slow-start technique to transmit initial packets in the windows of 1, 2, 4, 8, 16, and 32 will be $\Delta + (\Delta+\delta) + (\Delta+3\delta) + (\Delta+7\delta) + (\Delta+15\delta) + (\Delta+31\delta)$ where Δ is RTT and δ is the time to transmit consecutive packet in a window. Similarly, the time required for transmitting packets in the remaining windows can be estimated. Therefore, the total time required by the slow-start and AIMD technique will be $T_E = 27\Delta + 997\delta$.

In the following subsection we discuss packet overhead gain and time gain for the proposed Unexpected Packet based Congestion Control technique in congestion free network.

b) Proposed UPCC technique in congestion free network

In congestion free network, the proposed UPCC technique advocates the use of rwnd (receiver advertised window) as a window size for the transmission. Therefore, the number of ACKs required for transmitting the application of 1024 packets will be

$$X_P = \left\lceil \frac{Number\ of\ packets\ to\ be\ transfered}{rwnd} \right\rceil$$

Hence, X_P=21, implying 21 acknowledgements are required as compared to 27 acknowledgements in the existing technique. Thus, approximately 22% reduction in the number of acknowledgements is achieved through proposed UPCC technique.

Further, the time required to transmit an application of 1024 packets will be $T_P = (\Delta+49\delta) \times 20 + (\Delta + 23\delta) = 21\Delta + 1003\delta$. Thus, the proposed UPCC technique reduces the time by approximately 12% leading to lesser chance of congestion in the network.

In the following subsection we discuss packet overhead gain and time gainin congested network where packet loss may occur while transmitting this application.

c) Existing SS-AIMD [8] technique with single packet loss

Consider that 240th packet is lost while transmitting 11th window where 44 packets can be transmitted without waiting for the acknowledgement. When receiver receives out of turn packets, it sends duplicate acknowledgement. When the sender receives 3 duplicate ACKs, it indicates mild congestion. The existing algorithm updates ssthreshold, cwnd and window size as $ssthreshold = \text{cwnd}/2$, $cwnd = ssthreshold$, and $window\ size = \min(cwnd, rwnd)$ respectively. It retransmits the lost packet and enters in to AIMD phase directly. In AIMD phase, the window starts increasing additively from new calculated ssthreshold to rwnd(receiver advertised window) as arithmetic progression. Thus, the variation in window size will be 1, 2, 4, 8, 16, 32,40, 41, 42, 43, 44(loss occurs),22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43 and 37. Thus, $X_E = 34 + 3$, i.e., 37 acknowledgements which include three duplicate acknowledgements must be sent by the receiver to acknowledge the correct receipt of each window, causing the packet overhead on the network. The total time required for this process will be $T_E = 34\Delta + 991\delta$.

The subsection discusses the performance of the proposed technique under the same scenario.

d) Proposed UPCC technique with single packet loss

When 240th packet is lost and receiver receives the 241st packet after the receipt of 239th packet. It sends an acknowledgement for the correct receipt of 239th and expecting 240th packet, i.e., ACK (240). However, when it receives the 242nd packet it will resend the ACK (240). When sender receives two ACK (240), it retransmits the lost packet and keeps the $window\ size = rwnd$(receiver advertised window) indicating mild congestion. Thus, the proposed UPCC technique in congested network with single packet loss requires,$X_P = 23$, acknowledgements where 2 extra acknowledgements are used for informing loss of 240th packet. Approximately 38% and 24% reduction in the packet overhead and time is achieved respectively.

Further, the performance of the proposed UPCC technique is evaluated with respect to the existing SS-AIMD technique under severe congestion wherein the ACK is not received within the stipulated RTO time, i.e., when RTO timer expires.

e) Existing SS-AIMD technique [5] when RTO timer expires

Whenever sender's RTO timer expires before receiving acknowledgement, it indicates severe congestion. The sender presumes that the entire window is lost and starts retransmission by reducing the window size back to one. In the above example, while transmitting an application consisting of 1024 packets if the RTO timer expireswhen the sender window is 44. The algorithm updates ssthreshold, cwnd and window size as $\text{ssthreshold} = \text{current window}/2$, $\text{cwnd} = 1$ and $\text{window size} = \min(\text{cwnd}, \text{rwnd})$ respectively. It then starts retransmission by entering into slow-start phase where the window size increases exponentially from 1 up to new ssthreshold of 22 packets as geometric progression. After this it enters in AIMD phase, where the window size increases additively from new ssthreshold calculated to rwnd(receiver advertised window) as arithmetic progression. Thus, the variation in window size will be 1, 2, 4, 8, 16, 32, 40, 41, 42, 43, 44(timer expires), 1, 2, 4, 8, 16, 22, 23, 24, 25,…, 44 and finally 5. Hence, $X_E = 40$ number of acknowledgements must be sent by the receiver. The total time required by the existing SS-AIMD technique when timer expires will be $T_E = 44\Delta + 1028\delta$.

The performance of the proposed Unexpected Packet based Congestion Control (UPCC) technique when RTO timer expires is as follows:

f) Proposed UPCC technique when RTO timer expires

Whenever sender's RTO timer expires before receiving an acknowledgement, it indicates severe congestion due to single or multiple packet loss. In the proposed technique the sender starts retransmission of the first unacknowledged packet and waits for RTO time again. In its response the receiver will send the ACK of the last correctly received packet with the rwnd. This ACK will indicate that how many packets are lost. If one packet loss is perceived then the proposed technique will assume that the network had mild congestion and has recovered from it so it will continue with the rwnd received i.e., 50, 50, 50, 50, 50(timer expires), 1, 50, 50,…, 50, and 25. However, if multiple packets are lost then it updates rwnd and window size as $rwnd = rwnd/2$, and $window\ size = rwnd$ respectively. This reduction in window size will continue if repeatedly multiple packet loss occur , however, the window size will be boosted on successful transmission of a complete window as 50, 50, 50, 50, 50(timer expires), 1, 25, 50, 50, 50, …,50 and 48. Therefore, approximately 42% and 40% reduction in packet overhead is received in case of single and multiple packet loss when timer expires respectively. Further, reduction in the transmission time is perceived as approximately 35% and 32% lower for the proposed UPCC technique in the case of single and multiple packet lossrespectively when timer expires.

The above example demonstrated that as more and more packet are lost the performance of the proposed UPCC technique improves both in terms of packet overhead gain and time gain.

IV. Simulation Results

We perform extensive network simulations with the help of ns-2, the widely used open-source network simulator [20]. We compared our proposed Unexpected Packet based Congestion Control (UPCC) technique with traditional slow-start and AIMD technique (NewReno[12]) and found that proposed UPCC technique reduces the packet overhead by 22% to 40% as shown in Figure 12 and also reduces the time to transmit an application by 12% to 32% as depicted in Figure 13. The variations in packet overhead and time depend on the level of congestion present in the network. The simulations were conducted in three different categories as 1) congestion free 2) single packet loss and 3) multiple packet loss. Figures 14 and 15 gives the results for congestion free network that shows that proposed UPCC technique reduces packet overhead and time thus minimizing the chance of congestion in the network.

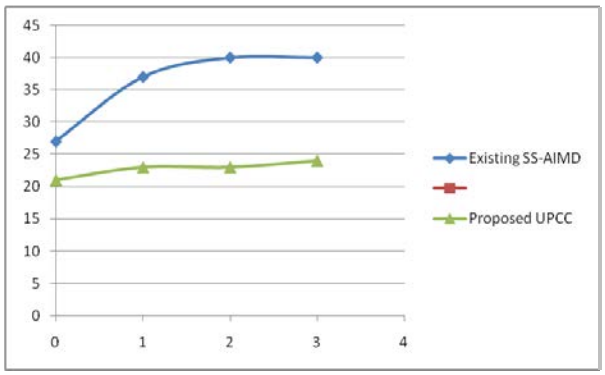


Figure 12 : Packet overhead gain of proposed UPCC vs SS-AIMD

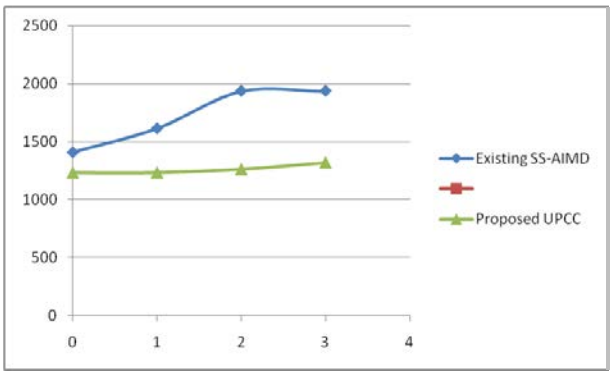


Figure 13 : Time gain of proposed UPCC vs SS-AIMD

Similarly, we conducted simulations for varying application sizes in case of multiple packet loss as shown in Figures 16 and 17 that clearly demonstrate that our proposed UPCC technique reduces packet overhead and time thereby minimizing the chance of congestion in the network.

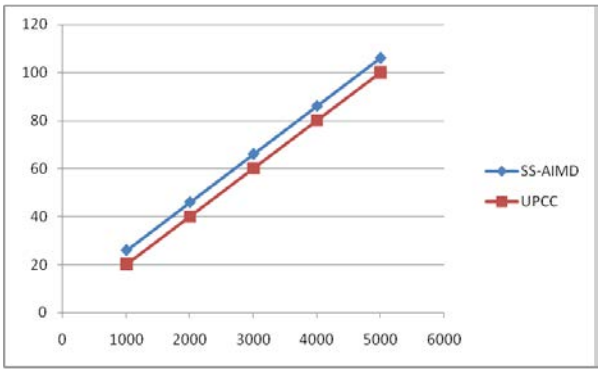


Figure 14 : Packet overhead gain of proposed UPCC vs SS-AIMD in congestion free network

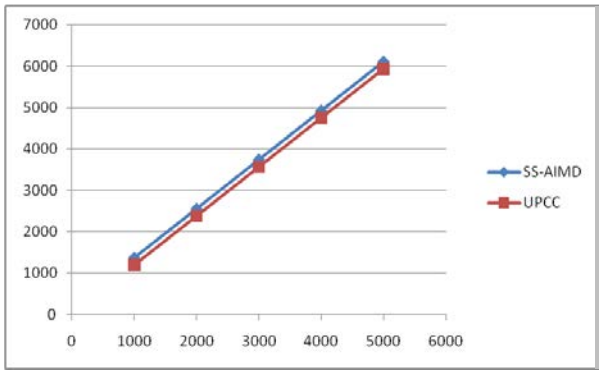


Figure 15 : Time gain of proposed UPCC vs SS-AIMD in congestion free network

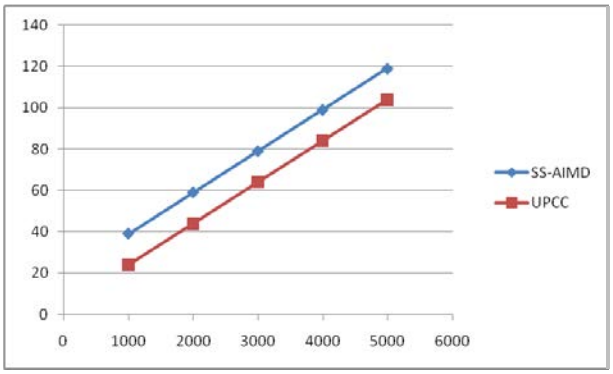


Figure 16 : Packet overhead gain of proposed UPCC vs SS-AIMD in congested network with multiple packet loss

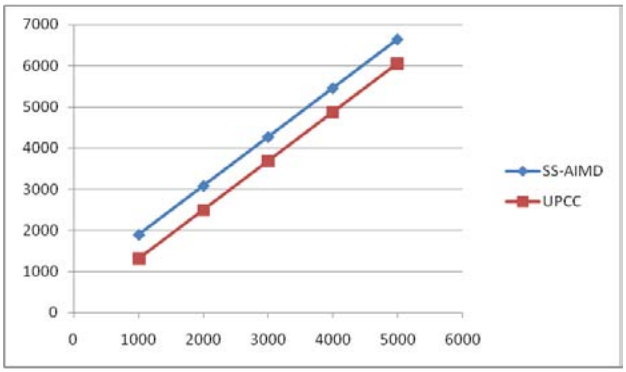


Figure 17 : Time gain of proposed UPCC vs SS-AIMD in congested network with multiple packet loss

V. Conclusions

In this paper we have demonstrated the benefit of using Unexpected Packet based Congestion Control (UPCC) technique. The simulation results shows that UPCC technique reduces the packet overhead and also reduces the time to transmit an application of various sizes as compared to the existing slow-start and AIMD technique.

References Références Referencias

1. R. Pan, P. Natarajan, C. Piglione, M. Prabhu, V. Subramanian, F. Baker, B. V. Steeg "PIE: A Lightweight Control Scheme To Address the Bufferbloat Problem," Cisco Systems, December 10, 2012, Internet Draft, draft-pan-tsvwg-pie-00 (work in progress).
2. Michael Welzl, "Network Congestion Control, Managing Internet Traffic," John Wiley & Sons Ltd, 2005.
3. Filipe Abrantes, and Manuel Ricardo, "On Congestion Control for Interactive Real-time Applications in Dynamic Heterogeneous 4G Networks," March 17, 2005. This work was funded by the Portuguese Science and Technology Foundation.
4. S. Floyd and K. Fall, Promoting the Use of End-to-End Congestion Control in the Internet, Networking, IEEE/ACM Transactions on, August 1999.
5. Jacobson, "Congestion avoidance and control," ACM Computer Communication Review, vol. 18, pp. 314–329, August 1988.
6. Srinivas Shakkottai and R. Srikant, "Network Optimization and Control, Foundations and Trendsin Networking" Vol. 2, No. 3 (2007) 271–379, 2008 S. DOI: 10.1561/1300000007.
7. Steven H. Low, Fernando Paganini, and John C. Doyle, Internet Congestion Control, IEEE Control Systems Magazine, February 2002.
8. M. Allman, V. Paxson, E. Blanton "TCP Congestion Control," Purdue University September 2009, Network Working Group, Request for Comments: 5681, Category: Standards Track.
9. Y. Cheng "Seeding RTO with RTT sampled during three-way handshake," Google. IncInternet Draft, draft-ycheng-tcpm-rtosynrtt-00.txt (work in progress), Intended status: Standard Updates: 3390, 2988, Creation date: June 30, 2010, Expiration date: January 2011.
10. B. Constantine, G. Forget, R. Geib, R. Schrage "Framework for TCP Throughput Testing," Internet Engineering Task Force (IETF), Request for Comments: 6349, ISSN: 2070-1721.
11. B. A. Forouzan "TCP/IP Protocol Suite," Tata McGraw-Hill, 3rd edition.
12. S. Floyd, T. Henderson, A. Gurtov "The New Reno Modification to TCP's Fast Recovery Algorithm," April 2004, Network Working Group, Request for Comments: 3782, Category: Standards Track.
13. V. Paxson, M. Allman, J. Chu, M. Sargent "Computing TCP's Retransmission Timer," June 2011, Internet Engineering Task Force (IETF), Request for Comments: 6298, Category: Standards Track, ISSN: 2070-1721.
14. RFC: 793, "TRANSMISSION CONTROL PROTOCOL," Defense Advanced Research Projects Agency, September 1981, University of Southern California.
15. Postel, J., "User Datagram Protocol," RFC 768, ISI, 28 August 1980.

16. Postel, J., "Internet Protocol," RFC 760, USC/Information Sciences Institute, January 1980.
17. RFC: 791, "INTERNET PROTOCOL," Defense Advanced Research Projects Agency, September 1981, University of Southern California.
18. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson "RTP: A Transport Protocol for Real-Time Applications," January 1996, Network Working Group, Request for Comments: 1889, Category: Standards Track.
19. Shudong Jin, Liang Guo, Ibrahim Matta, Azer Bestavros "A Spectrum of TCP-friendly Window-based Congestion Control Algorithms," Boston University, July 2002, This work was supported in part by NSF grants CAREERANI-0096045, ANI-0095988, ANI-9986397, and ITR ANI-0205294, and bygrants from IBM, Sprint Labs, and Motorola Labs.
20. Network Simulator ns-2. http://www.isi.edu/nsnam/ns/.

Global Journal of Computer Science and Technology: E Network, Web & Security
Volume 14 Issue 1 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Survey: Detection and Prevention of Wormhole Attack in Wireless Sensor Networks

By Akanksha Gupta & Anuj K. Gupta
RIMT College of Engineering & Technology, India

Abstract- Wireless Sensor Networks refers to a multi-hop packet based network that contains a set of mobile sensor nodes. Every node is free to travel separately on any route and can modify its links to other nodes. Therefore, the network is self organizing and adaptive networks which repeatedly changes its topology. The relations among nodes are restricted to their communication range, and teamwork with intermediate nodes is necessary for nodes to forward the packets to other sensor nodes beyond their communication range. The network's broadcasting character and transmission medium help the attacker to interrupt network. An attacker can transform the routing protocol and interrupt the network operations through mechanisms such as selective forwarding, packet drops, and data fabrication. One of the serious routing-disruption attacks is Wormhole Attack. The main emphasis of this paper is to study wormhole attack, its detection method and the different techniques to prevent the network from these attack.

Keywords: *wormhole attack, classification, detection mechanism, wsn, security, routing protocols*

GJCST-E Classification : *C.2.1*

Strictly as per the compliance and regulations of:

A Survey: Detection and Prevention of Wormhole Attack in Wireless Sensor Networks

Akanksha Gupta [α] & Anuj K. Gupta [σ]

Abstract- Wireless Sensor Networks refers to a multi-hop packet based network that contains a set of mobile sensor nodes. Every node is free to travel separately on any route and can modify its links to other nodes. Therefore, the network is self organizing and adaptive networks which repeatedly changes its topology. The relations among nodes are restricted to their communication range, and teamwork with intermediate nodes is necessary for nodes to forward the packets to other sensor nodes beyond their communication range. The network's broadcasting character and transmission medium help the attacker to interrupt network. An attacker can transform the routing protocol and interrupt the network operations through mechanisms such as selective forwarding, packet drops, and data fabrication. One of the serious routing-disruption attacks is Wormhole Attack. The main emphasis of this paper is to study wormhole attack, its detection method and the different techniques to prevent the network from these attack.

Keywords: wormhole attack, classification, detection mechanism, wsn, security, routing protocols.

I. Introduction

In Wireless Sensor Networks, the nodes use the open air medium to communicate with each other, in doing so they face sensitive security problems as compared to the wired networks. One such dangerous problem is wormhole attack. In this attack, two distant malicious nodes can plan together using either wired connection or directional antenna, to give an feeling that they are only one hop away. Wormhole attack can be executed in hidden or in sharing mode. Wormholes can either be used to examine the traffic throughout the network or to crash packets selectively or totally to affect the flow of information. The security mechanisms that are used for wired systems such as authentication and encryption are useless under hidden mode of wormhole attack because the nodes do not modify their headers but only forward these packets. But the attack in participating mode is more complicated, because if it once launched, it is difficult to detect.

WSN faces some challenges which are as follows:

1. Power Consumption – conservation of power is necessary and detection of some power saving routing protocol.
2. Multicast Routing – scheming of multicast routing protocol for a frequently changing WSN surroundings
3. Internetworking – Communication among wired system and WSN while maintaining synchronization.

II. Security Goals Designed for Wireless Sensor Networks

Security goals for WSN can be categorized as primary and secondary goals [35]. Some of the primary goals are Data Confidentiality, Data Authentication, Data Availability and Data Integrity and secondary goals are Data Freshness, Secure Localization, Self- Organization and Time Synchronization. The primary goals are also known as standard security goals.

Primary goals are as follows:

a) Data Confidentiality

Confidentiality is the capability to hide messages from a passive attacker such that every message communicated using the sensor network remains confidential. It is the most important concern in network security. A sensor node should not expose its data to its neighbors.

b) Data Authentication

Authentication ensures the consistency of the message by identifying its foundation. Attackers in sensor networks can not only responsible for the alteration of packets but can also insert additional fake packets [34]. Basically data authentication is used for the verification of the identity of the senders and receivers. Symmetric or Asymmetric mechanisms are used for data authentication in which sending and receiving nodes share secret keys. Because of wireless medium and unattended nature of sensor networks, it is very demanding to ensure authentication.

c) Data Integrity

Data integrity in wireless networks is desired to ensure the consistency of data and to verify that a message has not been altered, tampered with or changed. Though the system has secrecy measures, but still there is a possibility of alterations. The integrity of the system will be in dilemma when:

- A wicked node present in the network adds false data.
- Due to wireless channel unstable conditions can cause harm or loss of data [33].

Author α: RIMT College of Engineering & Technology, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India.
e-mail: gupta.akanksha75@gmail.com
Author σ: Head of the Department (Computer Science and Engineering), RIMT College of Engineering & Technology, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India.
e-mail: anujgupta@rimt.ac.in

d) Data Availability

Availability ensures whether the resources are free to be used by a node and whether the network is existing for the messages for communication. However, failure of the base station or cluster leader's availability will eventually threaten the entire sensor network. Thus data availability has a main importance for maintaining an operational network.

Secondary goals are as follows:

e) Data Freshness

Even though confidentiality and data integrity are guaranteed, there is also a need to make sure the freshness of each message. Basically, data freshness [33] ensures that the data is new and no old data have been replayed. To resolve this trouble related counter will be added into the packet to guarantee data freshness.

f) Self-Organization

A wireless sensor network is a usually an ad hoc network, in which every sensor node is independent and flexible such that each nodes is self-organizing and self-healing to different situations. No permanent infrastructure is present in a wireless sensor network for network management. This natural feature challenges the wireless sensor network security. So if self-organization is absent in a sensor network, then the harm that results from an attack or from the risky environment may be disturbing.

g) Time Synchronization

Most of wireless sensor network applications are based on some type of time synchronization. Moreover, sensors tries to calculate the end-to-end delay of a packet as it travels from source to destination sensor or node. A shared sensor network can require group synchronization [33] for purpose of tracking applications.

h) Secure Localization

The effectiveness of a sensor network is based on its ability to locate each sensor node in the network correctly and automatically. Now a days, sensor networks designed to locate faulty nodes which will require the accurate location information. An attacker can easily operate all the non protected location information by exposing the replaying signals and false signal strengths etc.

III. Wormhole Attacks

Wormhole attack contains two nodes that are connected to one another with the a medium that is not offered to normal nodes, due to which the nodes can communicate with one another over a range in which normal nodes cannot. These two colluding nodes are operated such that they shown like a neighbors to all the other nodes. In [Figure 1], suppose node-1 wants to send any data to node-25 through the network, so node-1 broadcasts the route request. Let node-Xs and node-Xd are the two colluder nodes in the locality of source node and destination node. Now Xs along with other nodes in the network gets the route request from source node, it replays the same request to Xd, Xd receives the request and de-capsulate it and rebroadcasts it to its neighborhood. After receiving the route request through Xd the destination node-25 will think that they are direct neighbors to source node-1, and it will reply to that route request. Xd will then capture that reply and using the same process it will send it to Xs; which further send to node-1. Thus node-1 and node-25 will believe that they are 2- hop neighbors. And complete communication will pass through Xs and Xd. This is one type of wormhole attack; many more number of variants are defined in the literature [3], [4], [5].

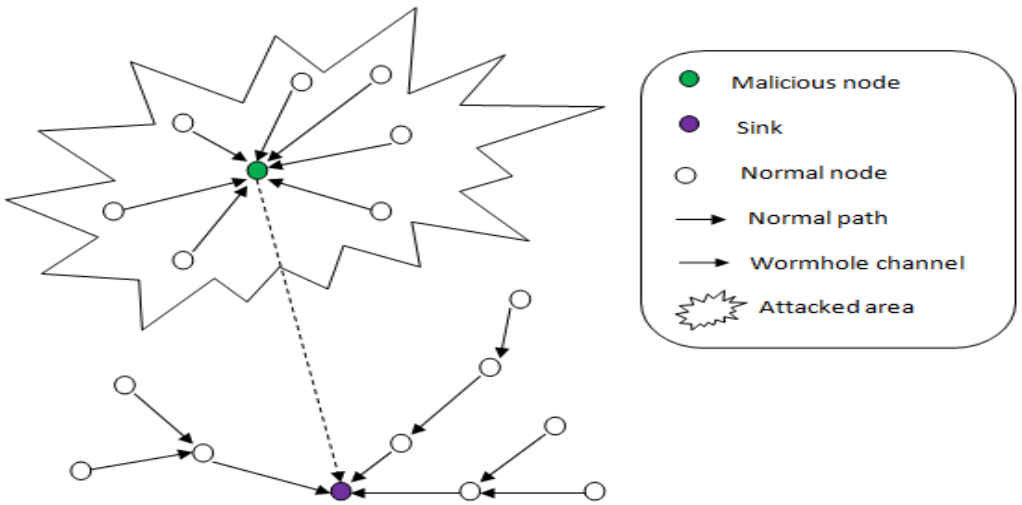


Figure 1

IV. Classification of Wormhole Attacks

In [6], [7], [8]; Wormhole attacks can be classified on the basis of:

1) Its Implementation
2) The medium used
3) The attackers
4) The location of victim nodes.

a) Classification based upon Implementation

This is the most important classification; which depends upon the behavior the attack is launched.

i. *Using Encapsulation*

In this manner, there are some nodes are occupied along the path (these nodes may or may not be conscious of wormhole) between xs and xd. The packet gets encapsulated at xs and travels through the path in encapsulated form to avoid the increase in the hop count. In this case the attackers are not directly connected to one another rather make the other nodes believe that they are directly connected. These packets are transmitted between xs and xd using a virtual tunnel. Once this attack is successfully launched, then all the paths will contain a link that will contain of link between xs and xd.

ii. *Using Out-Of-Band Channel*

These colluder nodes get connected directly through a out of band channel having high bandwidth. The channel can be obtained by a wired connection or using a wireless connections. The requirement of extra hardware made it difficult to launch, but provides a simplicity because it will not require any encapsulation/de-capsulation while the colluders are directly connected.

iii. *Using High Power Transmission*

This type of wormhole particularly launched from two colluder nodes that facilitates high power transmission potential.

iv. *Via Protocol Deviations*

In this case the attackers generate the wormhole by not following the protocol set of laws e.g. Some protocols suppose the nodes to wait for a while before retransmitting but the attackers keeps on broadcasting and do not obey this rule and thus trying to reach first at the destination and thus avoiding any future genuine requests to reach destination. If the future requests arrive at destination, they will be dropped, since a request passing through the colluder has previously been received.

b) Classification based upon Medium Used

On the basis of medium used, wormhole attacks can be classified as in-band and out-of-band wormhole attacks.

i. *In-Band Wormhole*

Same medium will be used by the attackers for creating link between them e.g. protocol deviations, packet relay and, encapsulation.

ii. *Out-Of-Band Wormhole*

Like normal network nodes attackers do not use the same medium, e.g. High Transmission Mode and Out-Of-Band Channel.

c) Classification based upon Attackers

i. *Self-Sufficient*

Here colluder nodes present themselves as normal nodes and thus all paths passes through them e.g. using high power transmission or out-of-band channel.

ii. *Extended Wormhole*

The colluder nodes extends the attacks beyond themselves to normal nodes and are unseen by themselves e.g. packet relay or encapsulation.

d) Classification based upon location of Victim nodes

i. *Simplex*

The victim node is present inside the range of only one attacker.

ii. *Duplex*

The victim node is present inside the range of both the attackers.

V. Literature Review

A significant amount of work have been prepared for the detection of wormhole attacks and the attackers. The work ranges from suggestion of extra and exclusive hardware to minor modifications in the system protocols and suggestion of smart ways of avoiding or detecting the wormholes. However some can need extra hardware and other require extra processing and battery life. This section shows a small review of the approaches proposed till date.

Hu et al. [16] proposed the method in 2003 based upon geographical and temporal packet leaches. In this method to avoid the wormhole, the geographical location or temporal location is used to bound the distance travelled by the packet. This approach is restricted by condition of GPS technology or the time synchronization. Lazos et al. [17] proposed a method in 2005 where a few nodes are mandatory to be equipped by GPS locators and directional antennas. This procedure uses “local broadcast keys” for safe communication between one another.

Tran et al. and Phuong et al. proposed TTM (Transmission Time based Mechanism) in 2007, where every node in the pathway work together and attack is identified through route setup stage by calculating transmission time among two nodes. Venkataraman et al. in 2009 proposed a graph theoretic mechanism for

the finding of wormhole attacks, which is right for proactive protocols.

Chen et al. [18] proposed a secure localization approach in 2010 based on the inconsistent set based resistant localization. Graaf et al. [19] proposed a dispersed detection approach based upon ranges of nodes for the detection of wormhole attacks. A Vani et al. [20] proposed a solution in 2011 that combines the decision anomaly, neighbor list count and hop count methods for AODV protocol. This procedure depends upon hierarchical processing of nodes and their respective neighbors. They used the hop count information available in the routing table of the nodes which needs that we need to store two copies of routing table of every node so as to maintain the track of earlier hop counts.

VI. Routing Protocols and Wormhole Attack

Various routing protocols are existing for WSN. Some of the often used routing protocols are considered in this section and the risk of wormhole attacks to such protocols is described. These routing protocols are classified into two types: proactive / table-driven protocols and reactive / demand-driven protocols [1]. AODV, DSR and Ariadne are reactive routing protocols and OLSR, DSDV and SEAD are proactive routing protocols.

a) OLSR (Optimized Link State Routing)

It is a proactive routing protocol in which information of the topologies get exchanged periodically. Hello messages are transmit to determine single hop neighbors. To allocate signaling traffic, flooding system is use. In this system each node forwards flooded message that was not forwarded by them earlier. The topology messages contains all the information about link states that are sent to all other nodes. With the help of this information, partial topology graph are obtained by every node after calculating the shortest path using symmetric relations. Now this system is open to wormhole attack [9] – [11]. Isolated nodes can send hello and topology manage messages are available at its colluding nodes to its personal neighbors for broadcasting fake information into the system. This will create two distant nodes to mistakenly believe themselves as neighbors, that leads to the failure of routing protocol.

b) DSDV (Destination Sequenced Distance Vector)

It is a proactive routing protocol, in which all the metric, destination routes, sequence number generated by the destination node and next hop to each destination are maintained in a table [1], [2]. Every node in the network acts as a router and the table gets updated periodically by exchange of messages among neighboring routers. This protocol is open to wormhole attack [9]. By using a tunnel, the colluding nodes surpass message between two distant nodes, suppose X and Y which will results X and Y to consider themselves as neighbors and they will publicize a hop count of one among each other. As a result of this false information, if the alternative route has hop count more than one then all other authenticated nodes will aim to send the messages through X to destination Y.

c) DSR (Dynamic Source Routing)

It is a reactive routing protocol because it discovers the required routes only after it has packets to transmit to the destination. It wants source route maintenance because during the utilization of the route, it is necessary to check the operation of the path and to report the sender regarding the errors [2]. It is at risk to wormhole attack and denial of service attack at the destination [9]. This protocol ensures forwarding of just the first RREQ that it will received and will reject all other RREQ packets for the same route. This RREQ packet contains the intermediate nodes and the hop count information. The route then established is used to send data packets. As wormhole attack ensures a fast channel for forwarding messages, so as compared to other paths RREQ packet through them will arrived at destination faster. This will result in only the wormhole path to be discovered as the route to destination. The wormhole attacker discards the data packets totally or partially that results in denial of service attack at the destination.

d) SEAD (Secure Ad-hoc Distance Vector)

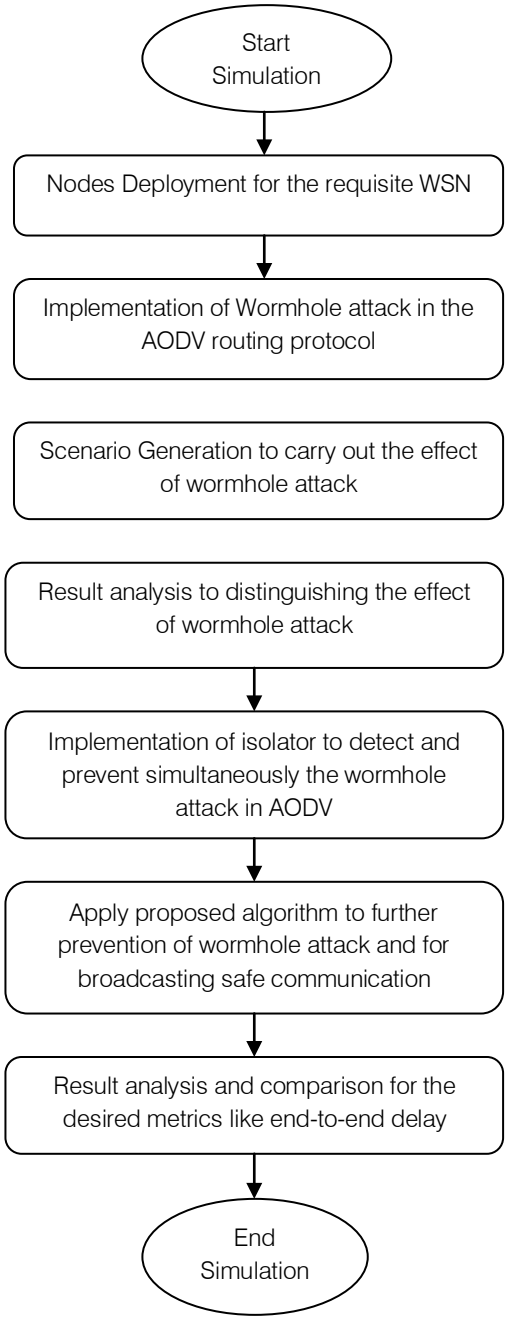
This protocol depends upon on one-way hash chains rather than asymmetric cryptograph and protects the network from uncoordinated attacks and DoS attacks. Several nodes have the ability to authenticate all other elements of the chain. This requires authenticating the metric of the routing table and the sequence number. The receiver should also verify the sender [2]. Thus, an enemy is not able to send routing message without compromising a node, as it does not give authentication code to its neighbors [12]. Although SEAD effectively handles replay attack, it is incapable to handle the wormhole attack [13] by a malicious node that are replaying the message from an unauthenticated node as a repeater.

e) AODV (Ad-hoc On-demand Distance Vector)

It is an on-demand routing protocol which broadcasts RREQ messages to its immediate neighbors for sending messages to final destination and in turn these neighbors rebroadcast them to their neighbors. This whole process continues unless until the RREQ message reaches the destination. On getting the initial RREQ message from the source, the destination node sends a RREP to the source node through the same reverse path [1], [15]. All the in-between nodes also put forward route entries in their respective table. The

neighboring nodes forward route error message to all its neighbors after detecting fault in any link to a node. This will again start a route discovery procedure to change the broken link. This AODV routing protocol is also at risk to wormhole attack [9].

As wormhole attack ensures a fast channel for forwarding messages, so as compared to other paths RREQ packet through them will arrived at destination faster. In this protocol, the destination rejects all the later on RREQ packets received, yet they are from authenticated node. Hence the destination chooses the fake path through wormhole for RREP [1].



The Prevention of Wormhole attack in AODV using WSN

f) Ariadne (A Secure On-Demand Routing Protocol for Ad-hoc Networks)

This protocol depends on symmetric cryptography and ensures that the source can authenticate each intermediate node in the route and the destination node authenticates the source. All intermediate node can eliminate or insert nodes in the list of nodes of the route request. It uses the key management protocol known as TESLA that focuses on the clock synchronization to authenticate routing

messages. TESLA uses per-hop hashing method [2]. An authentication done at each node does not only depends upon the information contained in the RREQ packet but also depends the authentication code of the preceding node. Ariadne protocol is free from over-flooding of RREQ attack because the attacker is prevented from replaying the message due to the network-wide shared secret key. It is necessary for each node to insert authentication code to every RREQ packet that it forwards. Then the source can be able to authenticate the origin of all individual data field in the RREP packet [14]. It is protected from rushing attack wormhole and attack [13] while successful route distortion requires RREQ to be tailored cautiously.

VII. Detection and Avoidance of Wormhole Attacks

From past few years the main area of research is the detection of wormhole attack. The most important task is to discover the occurrence of wormhole in the system [12], [24] – [31].

Detection of wormhole on the basis of the Hello control messages [26]. With the use of OLSR specifications, the percentage of HELLO Message Timing Intervals (HMTIs) which lies in a range enclosed by the amount of jitter. A range $R = [T - \delta, T + \delta]$ have been defined. If HMTI is in the range R, then it will considered to be valid; otherwise it is said to be out-of-protocol. A second check is made every time when the HMTI packet behavior is doubtful. On other side, a badly performing node would get coupled with it a comparatively large number of repeat packets, that would not be the case by an attacking node. In this way, the false positive alarms problem gets negotiated.

A new protocol known as Multi-path Hop-count Analysis (MHA) is proposed on the basis of hop-count analysis to stay away from wormhole attack [24]. It is supposed that very low or very high hop-count is not good for the network. The uniqueness of the hop-count analysis for detecting wormholes nodes is yet uncertain.

Wormholes nodes are detected by assuming that wormhole attacks have longer packet latency as compared to the normal wireless propagation latency in a single hop [10]. As the route during wormhole seems to be shorter, various new multi-hop routes be also channeled in the direction of the wormhole that leads to the longer queuing delays. The links having delays are considered to be doubtful links, as the delay might also takes place due to congestion as well as intra-nodal processing. The OLSR protocol is used for routing. This approach aims to sense the suspicious link and authenticate them in a two step process that is described below.

In first step, Hello packets has been sent to all the nodes that are within its transmission range. As soon as the receiver receives the Hello message, then it records the address of the sender and the time delay Δ left until it will be programmed to send its next Hello message. The node attaches the address of the sender and their respective values of time delay Δ that has recorded for piggybacked reply. When Hello reply is received by a node, then it checks for the information related to any of its outstanding requests. But if no such information is there, then it will suppose it as any other control packet. Otherwise, node checks the arrival time of Hello reply message to notice whether it is arrived within its scheduled timeout interval by considering the time delay Δ that occur at the receiver side. If the arrival time is within its timeout interval then link between itself and node is taken to be safe, otherwise doubtful and communication to that node is terminated by the sender until the verification process gets over.

In second step, a probing packet is sent to all the suspected nodes (that are detected in the previous step) by the sender.

If a suitable acknowledgement is received from any node X within its scheduled timeout interval then node X is considered to be safe. Otherwise the occurrence of wormhole is proved.

Both delay per hop indication (DelPHI) and hop count are monitored for wormhole detection [22]. The basic assumption is that the delay that packet experience in standard conditions for propagation of one hop will become too high under wormhole attack as the actual path between the nodes is shorter than the advertised path. This proposed methodology for wormhole detection is a two-step process.

The first phase has the route path information, gained from a set of disjoint paths from sender to receiver. Every sender will consist of a timestamp on a unique DREQ packet and send it to receiver after signing it. After receiving the packet for first time every node will adds its node ID then increase the hop count by 1 and rejects the packet next time onwards. After receiving each disjoint path the receiver send the DREP packets. This process is carried out for three times and the hop count and smallest delay information will be chosen for wormhole detection.

In second phase, the time difference between the packet it had sent to its neighbor and the reply received by it known as round trip time (RTT) is calculated. Delay per hop value (DPH) is evaluated as RTT/2h, where h stands for hop count to the particular neighbor. Under ordinary circumstances, a smaller h also have smaller RTT. However, smaller hop count will have larger RTT under wormhole attack. But one DPH value for node X exceed the consecutive one by several threshold, then path from node X to every another paths with DPH values greater than it is considered as under wormhole attack.

Similar propositions are made in SaW [29] and DaW [30]. In SaW, AODV protocol was used and in

DaW, DSR routing protocol was used. In these papers, security models have been planned and used to detect interruption. To detect the attacks, it will use statistical methods. If any link is identified to be doubtful, then existing information is used to detect the presence of a wormhole. In trusted model, nodes monitor their neighbors on the basis of packet drop pattern but not on the basis of number of drops. Other algorithm has been proposed in [30] to detect the presence of wormhole into system. In this algorithm, the source waits for RREP after sending the RREQ. The source receives a lot of RREP from different routes. By using the below expression we can find out the link with very high frequency:

$$Fi = pi / P, \text{ for all } Li$$
$$Fmax = max\ (Fi),$$

where r is the set of all obtained routes, *Li* is the ith link, pi is the number of times that *Li* appears in *r,* P is the total number of links in *r*, and *Fi* is the relative frequency that *Li* appears in *r*. If *Fmax* > *Fthreshold,* then check the information present in RREP of that route. The node will be malicious if the value of correlation coefficient for packets dropped is greater than the pre-set threshold value t, then it will inform the operator otherwise continue with routing process.

According to [29] and [30], the regular link frequency analysis may lead to fake detection of wormhole attacks. Though, these recognize the performance of a wormhole as they record the total number of packet drops rather than the pattern of drop.

The wormhole attack can be detected using multipath routing [27]. When a source node wants a new route, it will broadcast the RREQ into the network and wait for responses. Then the in-between node will forward only the first RREQ packet. After receiving the first RREQ the destination will wait for a while to gather all the obtained routes. A new scheme known as Statistical Analysis of Multi-path (SAM) is projected in [27]. SAM uses *Pmax* (i.e. maximum probability of relative frequency of a link to occur in the set of all obtained routes from one route discovery) and Ø (i.e. difference between the most frequently appeared link and the second most frequently appeared links in the set of all obtained routes from one route discovery), which will be higher in the presence of wormhole attack. Relative Frequency is calculated using probability mass function (PMF) which is more for a network that is under wormhole attack as compared to a normal network. The performance of Dynamic Source Routing (DSR) and On-demand multipath routing (MR) protocol are compared under wormhole attack.

A cluster based counter-measure known as WHIDS [28] is proposed for the wormhole attack. By using MATLAB simulation results the effectiveness of WHIDS are revealed for detecting wormhole attack. This method, yet not been experienced in existence of multiple wormhole attacks.

Vu et al. proposed the technique to detect the existence of wormhole node using two phases [31] as in [10] and [22]. The first phase contains of two methods: In first method, the computation of round-trip-time (RTT) among the source and all of its immediate neighbors is measured. In second method, source identifies its one-hop and two-hop neighbors to form its neighbor set. If it is originated that the destination is not the neighbor of source node then the link between them will be taken as suspicion. After detecting the doubtful links, the next phase is used to verify the presence of wormholes for exchange of messages by using the RTS / CTS mechanism.

Table 1 represents multi-aspect comparison among eight different wormhole detection techniques. Significant aspects like false alarm detection, the node mobility along with QoS parameters are considered for each detection technique. This qualitative study have been supported by quantitative one also for several algorithms using the network simulator tool.

Table 1 : Summary about the detection methods for wormhole attack.

Method	Mobility	QoS Parameter	Synchronization	False detection
WORMEROS [31]	Topological change is not considered	Not considered	Time synchronization not required. RTT between source node and destination node is considered	Both false positive and false negative alarms are considered
HMTIs [26]	Handled weakly. Topologically robust, short range worm-hole can be detected	Jitter and delay	Not required. Since PSD profiling is done locally	Used PSD to detect false positive alarm
Farid et al. [10]	Not considered	Packet processing time, queue	Some time delay added to detect suspicious links	Not handled
DelPHI [22]	Not considered	Delay	Not required	Not handled

SAM [27]	Cluster and uniform topology considered	Not considered	Not considered	Not handled
SaW [29]	Not considered	Not considered	Not considered	Failed to detect
DaW [30]	Not considered	Delay parameter	Not considered	Failed to detect
WAP [23]	Maximum transmission distance is calculated	Delay per hop	Only the source node is synchronized	Not handled

References Références Referencias

1. R.H. Khokhar, Md. A. Ngadi, S. Manda. "A Review of Current Routing Attacks in Mobile Ad Hoc Networks", International Journal of Computer Science and Security, 2008.
2. D. Vivian, E.A.P. Alchieri, "Evaluation of QoS Metrics in Ad Hoc Networks with the use of Secure Routing Protocols", In Network Operations and Management Symposium, 2006.
3. Meghdadi M, Ozdemir S," A Survey of Wormhole-based Attacks and their Counter measures in Wireless Sensor Networks", IETE Tech Rev 2011.
4. S.Vijayalakshmi and S.Albert Rabara "Weeding Wormhole Attack in MANET Multicast Routing Using Two Novel Techniques - LP3 and NAWA2", International Journal of Computer Applications, February 2011.
5. Jhaveri, R.H.; S.J.; Jinwala, D.C.; "DoS Attacks in Mobile Ad Hoc Networks: A Survey," Advanced Computing & Communication Technologies (ACCT), 2012.
6. V. Mahajan, M. Natu "Analysis of wormhole intrusion attacks in MANETS", In IEEE Military Communications Conference (MILCOM), 2008.
7. H.S. Chiu and K. Lui. "Wormhole Detection Mechanism for Ad Hoc Wireless Networks". In Proceedings of International Symposium on Wireless Pervasive Computing, 2006.
8. R. Maulik, N. Chaki. "A Comprehensive Review on Wormhole Attacks in MANET". In 9th International Conference on Computer Information Systems and Industrial Management Applications, 2010.
9. Yih-Chun Hu, Adrian Perrig. "Packet leashes: A Defense against Wormhole Attacks in Wireless Ad Hoc Networks". In 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, 2003.
10. F. Nait-Abdesselam, B. Bensaou. "Detecting and Avoiding Wormhole Attacks in Wireless Ad hoc Networks", IEEE Communications Magazine, 2008.
11. John S. Baras, George Theodorakopoulos. "Intrusion Detection System Resiliency to Byzantine Attacks: The Case Study of Wormholes in OLSR". In IEEE Military Communications Conference, 2007.
12. Y.-C. Hu, A. Perrig, D.B. Johnson. "SEAD: Secure Efficient Distance Vector Routing for Mobile Wireless Ad Hoc Networks". In Proceedings of Fourth IEEE Workshop on Mobile Computing Systems and Applications, 2002.
13. Kamanshis Biswas, Md. Liakat Ali. "Security Threats in Mobile Ad Hoc Network". Paper submitted to the Department of Interaction and System Design, School of Engineering at Blekinge Institute of Technology, 2007.
14. Y.-C. Hu, D.B. Johnson. "Ariadne: A Secure On-Demand Routing Protocol for Ad Hoc Networks", Wireless Networks, 11(1-2), 2005.
15. Bounpadith Kannhavong, Hidehisa Nakayama, Abbas Jamalipour. "A Survey of Routing Attacks in Mobile Ad Hoc Networks", IEEE Wireless Communication, 2007.
16. Hu, Y.-C.; Perrig, A.; Johnson, D.B.; "Packet leashes: a defense against wormhole attacks in wireless networks,". Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies , April 2003.
17. Lazos, L.; Poovendran, R.; Syverson, P.; Chang, L.W.; , "Preventing wormhole attacks on wireless ad hoc networks: a graph theoretic approach," Wireless Communications and Networking Conference, March 2005.
18. Honglong Chen , Wei Lou , Zhi Wang, A secure localization approach against wormhole attacks using distance consistency, EURASIP Journal on Wireless Communications and Networking, April 2010.
19. R. Graaf, I. Hegazy, J. Horton. "Detection of wormhole attacks in wireless sensor networks," Springer book chapter Ad Hoc Networks, 2010.
20. A.Vani, D. Sreenivasa Rao, "A Simple Algorithm for Detection and Removal of Wormhole Attacks for Secure Routing In Ad Hoc Wireless Networks", International Journal on Computer Science and Engineering (IJCSE), June 2011.
21. V. Mahajan, M. Natu. "Analysis of wormhole intrusion attacks in MANETS". In IEEE Military Communications Conference, 2008.
22. H.S. Chiu and K. Lui. "DelPHI: Wormhole Detection Mechanism for Ad Hoc Wireless Networks". In Proceedings of International Symposium on Wireless Pervasive Computing, 2006.
23. S. Choi, D. Kim, J. Jung. "WAP: Wormhole Attack Prevention Algorithm in Mobile Ad Hoc Networks". In International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, 2008.

24. Shang-Ming Jen, Chi-Sung Laih. "A Hop-Count Analysis Scheme for Avoiding Wormhole Attacks in MANET", 2009.
25. D. Djenouri, O. Mahmoudi, M. Merabti, "On Securing MANET Routing Protocol Against Control Packet Dropping". In IEEE International Conference on Pervasive Services, 2007.
26. M.A. Gorlatova, P.C. Mason, L. Lamont, R. Liscano, "Detecting Wormhole Attacks in Mobile Ad Hoc Networks through Protocol Breaking and Packet Timing Analysis". In IEEE Military Communications Conference, 2006.
27. N. Song, L. Qian, X. Li. "Wormhole Attacks Detection in Wireless Ad Hoc Networks: A Statistical Analysis Approach". In Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium, 2005.
28. D.B. Roy, R. Chaki, N. Chaki. "A New Cluster-based Wormhole Intrusion Detection Algorithm for Mobile Ad-hoc Networks", 2009.
29. M.S. Sankaran, P.S. Das, S. Selvakumar. "A Novel Security model SaW: Security against Wormhole attack in Wireless Sensor Networks". In Proceedings of International Conference on PDCN, 2009.
30. Khin Sandar Win. "Analysis of Detecting Wormhole Attack in Wireless Networks", World Academy of Science, Engineering and Technology, 2008.
31. H. Vu, A. Kulkarni, K. Sarac, N. Mittal. "WORMEROS: A New Framework for Defending against Wormhole Attacks on Wireless Ad Hoc Networks". In Proceedings of International Conference on Wireless Algorithms Systems and Applications, 2008.
32. R. Maulik, N. Chaki. "A Comprehensive Review on Wormhole Attacks in MANET". In Proceedings of 9th International Conference on Computer Information Systems and Industrial Management Applications, 2010.
33. Ian F. Akykildiz, Yogesh Sankarasubramaniam and Erdal Cayirci, "A Survey on Sensor Networks", IEEE Communication Magazine, 2002.
34. Y. Wang, G. Attebury, and B. Ramamurthy, "A Survey of Security Issues in Wireless Sensor Networks," IEEE Communication Surveys Tutorials, 2006.
35. John Paul Walters, Zhengqiang Liang, Vipin Chaudhary, "Wireless Sensor Network Security: A Survey", Security in Distributed, Grid and Pervasive Computing Yang Xiao, 2006.

This page is intentionally left blank

Global Journal of Computer Science and Technology: E Network, Web & Security
Volume 14 Issue 1 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Survey: Hierarchal Routing Protocol in Wireless Sensor Networks

By Neha Saini & Jasmeet Singh
RIMT College of Engineering & Technology, India

Abstract- The wireless sensor networks (WSNs) has been grown immensely in the past few decades. Researcher had proposed a number of routing protocols for WSN. WSN has two type of architecture layered and cluster architecture. We classify various clustering approaches based on different criterion in section [3]. Hierarchical Clustering protocols discussed in section [4] have extensively been used to achieve network scalability, energy efficiency and network lifetime. In this paper we discuss the challenges in design of WSN, advantages and objectives of clustering, various clustering approaches. We present a detailed survey on proposed clustering routing protocol in WSN literature.

Keywords: *wireless sensor networks; clustering routing; cluster construction; data transmission; taxonomy.*

GJCST-E Classification : *C.2.1*

Strictly as per the compliance and regulations of:

A Survey: Hierarchal Routing Protocol in Wireless Sensor Networks

Neha Saini [α] & Jasmeet Singh [σ]

Abstract- The wireless sensor networks (WSNs) has been grown immensely in the past few decades. Researcher had proposed a number of routing protocols for WSN. WSN has two type of architecture layered and cluster architecture. We classify various clustering approaches based on different criterion in section [3]. Hierarchical Clustering protocols discussed in section [4] have extensively been used to achieve network scalability, energy efficiency and network lifetime. In this paper we discuss the challenges in design of WSN, advantages and objectives of clustering, various clustering approaches. We present a detailed survey on proposed clustering routing protocol in WSN literature.

Keywords: wireless sensor networks; clustering routing; cluster construction; data transmission; taxonomy.

I. Introduction

Wireless sensor network consist of tiny devices called sensor node and sink called base station. Sensor node sense and collect information from surrounding environment in which they lie and transfer it to sink. Various application areas such as security surveillance, military reconnaissance, habitat monitoring, medical and health, disaster management, industrial automation, etc make use of WSN to sense data in harsh environment. In above mentioned applications, reliability and on time delivery of sensory data is must for the critical mission success. Major challenges with wireless sensor networks are their limited source of energy, high traffic load and the coverage constraint. Routing of data in WSN has been one of the challenging areas for researchers [1].

In most wireless sensor network (WSN) applications network have the capability to operate unattended in harsh environments. Nodes in such environments are energy constrained and their batteries cannot be recharged .Such environment demands energy-aware routing and data aggregation protocols providing high scalability in order to maximize network lifetime.

Routing protocols in WSN, on the basis of network structure are categorized in to 3 main categories [2]:-

1. Flat
2. Hierarchical
3. Location based

In particular, hierarchical routing protocols (explained in section 4) offer significant savings in total energy utilization in WSN. In hierarchical routing protocols, sensor nodes organized the in to clusters. Each cluster is governed by a cluster-head and only heads send messages to a BS. Research community widely accepted the grouping of sensor node in cluster to achieve objectives such as scalability, prolonging network lifetime and high energy. Advantage of this method is it saves energy by data aggregation by CH. Less the energy utilization, the more the network life time in WSN. But this method of clustering may commence overhead due to the cluster organization and maintenance, but it has been verified that cluster-based protocols demonstrate better energy consumption and performance in comparison to flat network topologies for large-scale WSNs.

II. Related Work

Kumarawadu et al. [4] present a survey on clustering protocol and categorized them on the basis of CH selection and cluster formation parameters. In the survey author discuss the design issues and performance challenges in clustering protocol based on the taxonomy of neighbourhood information based clustering protocol, identity-based clustering approach, and biologically encouraged clustering algorithms and probabilistic clustering protocol.

Arboleda et al. [3] briefly discussed LEACH-based protocols, proactive and reactive protocol and presented a survey comparing various clustering protocols. Some concept such as clustering advantages, cluster types, cluster structure of clustering process explained in detail.

Jiang et al. [6] analyzed hierarchical routing protocol, author compare these protocols on eight parameter of clustering. Author highlight the three important advantages of clustering process for WSNs, such as more less overheads, scalability, and easy maintenance, and then present a categorization of WSN clustering schemes.

Deosarkar et al. [5] focus on CH election criteria based on three metrics deterministic, adaptive and combined metric. The author analyse the cost of CH

Author α : Research Scholar, Department of Computer Science and Engineering, RIMT College of Engineering & Technology, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India.
e-mail: neha13choudhary@gmail.com
Author σ : Assistant Professor, Department of Computer Science and Engineering, RIMT College of Engineering & Technology, Mandi Gobindgarh, Fatehgarh Sahib, Punjab, India.
e-mail: jasmeetgurm@gmail.com

election and evaluate it against cluster formation and distribution of CHs and concluded that there is a need of more scalable, energy capable and efficient clustering scheme in WSNs for data aggregation.

Deng [13] focuses on design issues and relative analysis of WSN clustering routing protocol for increasing the network lifetime. The authors analyzed numerous challenging issues that influence design of routing techniques in WSNs, and categorized routing algorithm with comparative analysis.

Xu et al. [11] consider six clustering protocol and compare them on various parameters such as data gathering robustness, scalability, network lifetime, security energy conservation.

Maimour et al. [7] considered nine distinctive clustering protocols based on two categories, pre-established clustering routing protocol and on-demand clustering routing protocol clustering routing protocols to attain energy conservation in WSNs and also discuss clustering protocol from the point of view of data routing.

III. Clustering

In order to study different clustering protocol we need to have knowledge of clustering parameters and its taxonomy. The objective of clustering protocols is to increase scalability, balance load, improve energy consumption, fault tolerance, efficient energy/resource, latency reduction, guarantee of connectivity and provide robustness in a WSN [3].

a) Challenges of clustering

Wireless Sensor Networks present vast challenges in terms of implementation. There are several key attributes that

Designers must carefully consider which are of particular importance in wireless sensor networks [3], [4].

- Cost of Clustering
- Election of Clusters and Cluster heads
- Real-Time Operation
- Cluster management (Synchronization)
- Data gathering
- Repair method
- QoS (Quality of Service)

b) Clustering parameter

Nodes and CH mobility:- Various published approaches assumed the sensor nodes to be stationary, such networks are stable thus it is easy to maintain intercluster and intracluster Communication [4]. But in case of sensor node mobility we need to re-elect the CH periodically and maintain cluster organisation continually.

Type of nodes: sensor nodes are of two types based on clustering approaches. In Homogenous network all sensor nodes have same functionality and in Heterogeneous network some sensor node are equipped with higher capabilities and complex hardware.

Cluster count: cluster count can be fix or variable depending up on which clustering technique is used [5]. In probabilistic and randomized approaches CH are not predetermined thus the cluster formation process result in to variable no. of clusters. Cluster count is fixed for approaches where the CH is predetermined.

Cluster-head election: Various published approaches adopt various criteria for selection of CH. The sensor node in every cluster elects a leader among all the node either on randomized basis or follow a probabilistic approach or based on some other criteria (such as based on residual energy, node degree etc.)

Cluster formation process: cluster formation technique are of two type centralized or distributed earlier approaches followed centralized or hybrid approach ,when CHs are just one or more coordinator nodes are used to partition the whole network off-line and control the cluster membership[6].But nowadays as time efficiency is important distributed approach is followed.

Communication among nodes: In clustering two type of communication can occur intercluster communication or intra cluster communication both can be further of two type single hop and multi hop. Earlier clustering approaches assume the communication among its nodes and CH to be single hop but nowadays various approaches are published which provide multihop communication in intracluster.

Overlapping: Overlapping in clustering is said to occur when a sensor node is shared by more than one cluster. Overlapping provide better routing efficiency and also fasten up cluster formation process [7]. Some published approaches allow overlapping, some try to have minimum overlap some not at all permit overlaps

c) Classification of clustering approaches

Clustering approaches varies depending on various features. On the basis of functionality and characteristics of sensor nodes in cluster clustering approaches are categorised in to two categories homogenous algorithm and heterogeneous algorithm. Heterogeneous sensor networks consist of two type sensors, common sensors (lower capabilities sensor, used to sense data) and sensor equipped with complex hardware (sensor with higher capabilities ,does the task of data aggregation etc). Homogeneous networks consist of sensor node with same characteristics, hardware and processing capabilities [8]. Based on cluster formation clustering approaches are of two type centralized and distributed algorithm.

Considering the network structure, there are two type of routing protocol in WSN: flat and hierarchical routing protocol [9]. In flat routing protocol all the node have same functionality. But effective for small scale network where as hierarchical routing protocol is

suitable for large scale network. Most hierarchical routing protocols are having cluster-based organization of nodes to imply data aggregation, thus saving significant amount of energy. In hierarchical network each cluster has a cluster head (CH) which performs the specialized task of data aggregation and fusion, and several sensor nodes act as members. The cluster formation process has two-level hierarchy where cluster head form higher level and member nodes in cluster form the lower level. The sensor nodes send data sense by them to their corresponding cluster-head periodically. CH nodes aggregate this data remove redundant information and transmit it to sink or base station (BS) directly or through multihop routing. However, most of the time CH nodes need to send data at higher distances than common member nodes of cluster, thus they spending greater energy than common nodes. Common solution to this problem is load balancing among the sensor nodes by re-electing new CHs periodically.

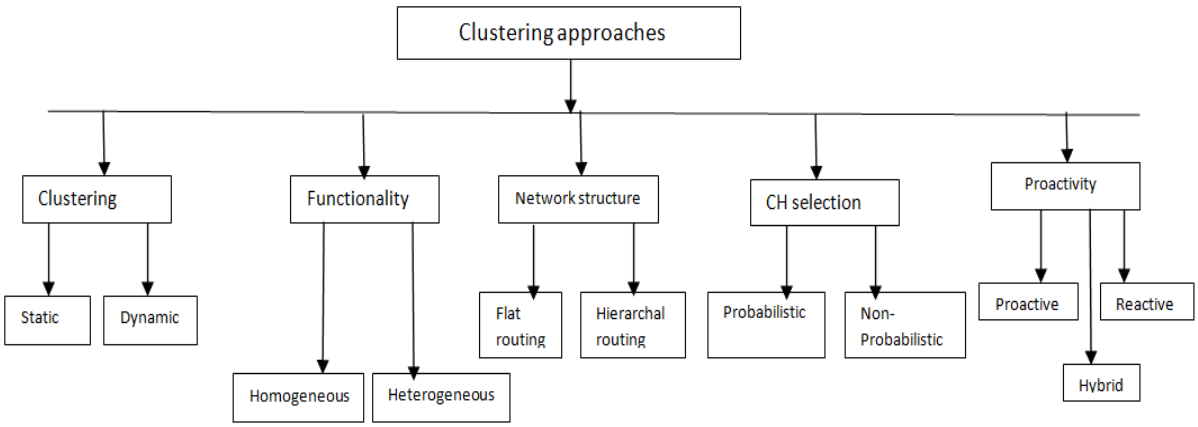


Figure 1 : classification for various clustering approaches

Another classification of clustering approaches is dynamic clustering and static clustering [11]. Cluster formation process is called dynamic when its CHS re-election is either event driven or periodic and react and adjust appropriately with cluster as well as network topology otherwise it is called static clustering approach .Dynamic clustering approach is very useful for sensors in WSN as it improve the network lifetime and manage the consumption of energy .

Most Clustering approaches are categorised in to two categories probabilistic and non probabilistic [12]. In Probabilistic approach a prior probability assigned to each sensor node is used to determine the initial CHs where as in non probabilistic clustering approach, a deterministic criteria for CH selection and cluster formation is followed and is based on nodes' proximity and on the data received from neighbouring nodes The typical clustering hierarchical protocols in WSNs include LEACH, EEHC, and HEED and their extensions. These are probabilistic algorithm some of them (LEACH and EEHC) follows random approach for CH election where as HEED is a hybrid approach where primary criteria followed by secondary criteria considered for CH.

Based on proactivity, clustering routing protocol can be categorized into proactive, reactive, and hybrid [10]. In proactive protocol, all routes between source and the BS are established before they are really needed in spite of data traffic. Once a message arrives, it go along a predestined route to the BS. Whereas, no predestined routes exist in reactive protocol, in which the routing is selected when a message desires to be delivered from source node to the BS. Hybrid approaches use a blend of the above two approaches. For this sort of clustering routing, occasionally proactive clustering mode is adopted, but at other period reactive mode is used.

IV. Hierarchical Routing Protocols

a) LEACH

Low-Energy Adaptive Clustering Hierarchy was first hierarchical routing clustering protocols. In WSNs, we need energy efficient network protocol such as LEACH is due to the fact that nodes in the WSNs are battery operated and have limited energy. In the LEACH protocol, the sensor nodes organize themselves into clusters each cluster is governed with cluster head (CH). Leach do load balancing by randomized alternation of cluster heads among all the sensor node in the network. This randomized approach is adopted to delay the first node death by distributing the load among all nodes in network. Cluster heads not only collect data from their clusters, but also aggregate the gathered data for reducing the data to be sent to the Base station, for less energy dissipation, to increase the network life time. Sensor nodes select themselves to be CHs at any time with some probability. The decision to nominate a node as cluster head is taken periodically. The elevation decision is to be ended only by each node free of other

nodes. This is done to reduce overhead in cluster head organization. The Threshold function is defined as [14]

$$T(n) = \frac{p}{1 - p(r \ mod \ (1/p))} \quad if \ n \in G$$

Where

n is the total node, r is the present round number, P is the probability of a node to be a cluster head and G is the set of nodes that have been not nominated as cluster heads in the previous 1/P rounds. Every node during cluster head selection will create a random number in between 0 and 1. The node will be converted into a cluster head if the number is less than the threshold (T (n)).

b) TEEN

TEEN stands for Threshold sensitive Energy Efficient sensor Network protocol (TEEN) [17]. It is a mixture of data-centric protocols and hierarchical clustering routing protocol and intended for real-time applications. It is a reactive protocol, quickly respond to sudden changes of some of the feature observed in the WSN (e.g., pressure).The protocol initially goes through cluster formation and cluster head selection. The CHs then transmit two thresholds to sensor nodes in their clusters. These are soft and hard thresholds for the sensed feature:

Hard Threshold (HT): It is the value, of the feature below which, the node sensing this value must turn on its transmitter and inform its cluster head.

Soft Threshold (ST): It stimulates the node to switch on its transmitter and inform the sensed data to its cluster head if change in the value is greater than the ST.

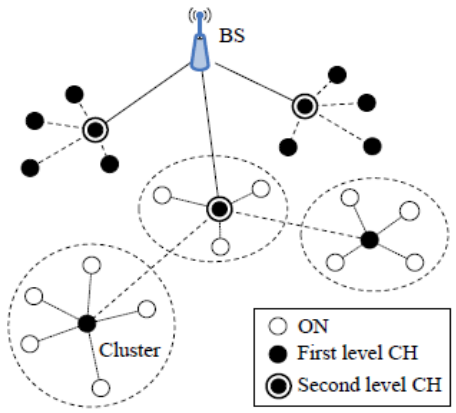


Figure 2 : Clustering topology in TEEN

A node will send data only when the sensed value is below the HT or alteration in the value is above the ST the. However, TEEN cannot be functional for sensor networks where sensor readings should be conveyed to the Sink in regular intervals, as the values of the feature may not accomplish the threshold at all. Furthermore, we have a number of shattered time-slots in TEEN protocol and there is forever likelihood that the sink may not be able to differentiate dead and alive nodes. Another drawback of the protocol is that the message broadcast is done by CHs only. If CHs are not in each other's transmission radius, the messages will be lost.

c) APTEEN

APTEEN-Adaptive Threshold sensitive Energy Efficient sensor Network protocol [18] is an expansion to TEEN protocol and goal at both having periodic data collections and reacting to real-time events. The structural design of APTEEN is same as in TEEN. In APTEEN firstly the clusters are formed by base station, the cluster heads relay the attributes, the transmission schedule and the threshold values to all nodes. Cluster heads achieve data aggregation in APTEEN in order to save energy. It supports three query types: one-time, to take a snapshot view of the network; historical, to analyze precedent data values; and determined to monitor an event at a time.

d) EECS

An Energy Efficient Clustering Scheme (EECS) [15] is a clustering algorithm in which cluster head candidates compete for the ability to elevate to cluster head for a given round. This competition involves candidates broadcasting their residual energy to neighbouring candidates. If a given node does not find a node with more residual energy, it becomes a cluster head. Cluster formation is different than that of LEACH. LEACH forms clusters based on the minimum distance of nodes to their corresponding cluster head. EECS extends this algorithm by dynamic sizing of clusters based on cluster distance from the base station. The result is an algorithm that addresses the problem that clusters at a greater range from the base station requires more energy for transmission than those that are closer. Ultimately, this improves the distribution of energy throughout the network, resulting in better resource usage and extended network life time. EECS is a LEACH-like clustering scheme, where the network is partitioned into a set of clusters with one cluster head in each cluster. Communication between cluster head and BS is direct (single-hop).In the network deployment phase, the BS broadcasts a "hello" message to all the nodes at a certain power level. By this way each node can compute the approximate distance to 0the BS based on the received signal strength. It helps nodes to select the proper power level to communicate with the BS. Also this distance is used to balance the load among cluster heads. In cluster head election phase, well distributed cluster heads are elected with a little control overhead. And In cluster formation phase, a

novel weighted function is introduced to form load balanced clusters.

e) HEED

Hybrid Energy-Efficient Distributed Clustering (or HEED) is energy efficient clustering protocol for wireless sensor networks, with a focus on efficient clustering by proper selection of cluster heads based on the physical distance between nodes. The main objectives of HEED are to [13]:

- Distribute energy consumption to prolong network lifetime;
- Minimize energy during the cluster-head selection phase;
- Minimize the control overhead of the network.

The most important aspect of HEED is the method of Cluster head selection. Cluster heads are determined based on two important parameters [13]:

1) The residual energy of each node is used to probabilistically choose the initial set of cluster heads. This parameter is commonly used in many other clustering schemes.
2) Intra-Cluster Communication Cost is used by nodes to determine the cluster to join. This is especially useful if a given node falls within the range of more than one

Cluster head. In HEED it is important to identify what the range of a node is in terms of its power levels as a given node will have multiple discrete transmission power levels. The power level used by a node for intra-cluster announcements and during clustering is referred to as cluster power level [13]. Low cluster power levels promote an increase in spatial reuse while high cluster power levels are required for intercluster communication as they span two or more cluster areas.

Therefore, when choosing a cluster, a node will communicate with the cluster head that yields the lowest intra-cluster communication cost. The intra-cluster communication cost is measured using the Average Minimum Reach ability Power (AMRP) measurement. The AMRP is the average of all minimum power levels required for each node within a cluster range R to communicate effectively with the cluster head i. The AMRP of a node i then become a measure of the expected intra-cluster communication energy if this node is elevated to cluster head. Utilizing AMRP as a second parameter in cluster head selection is more efficient then a node selecting the nearest cluster head [13].

f) PEGASIS

PEGASIS is a data-gathering and near-optimal chain-based algorithm. Power-Efficient Gathering in Sensor Information Systems [8] protocol reduces the consumption by creation of a chain structure containing of all nodes and simultaneously do data aggregation across the chain. According to PEGASIS algorithm if nodes made a chain from source to sink, among all the node across chain only one node will send the data to base station in a given transmission time-frame. Data-aggregation occurs at all node in the sensor network to pervade all important information across the network.

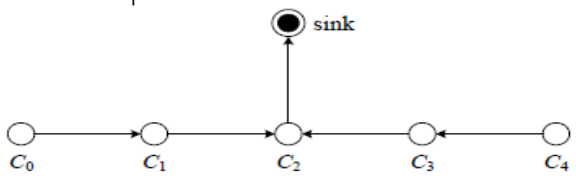


Figure 3 : Data transmission scheme in PEGASIS

In PEGASIS in spite of multiple nodes only one node in a chain transmit data to the BS. It increases the network life time, when all nodes take turns in communicating with the BS and node communicate only with their nearby neighbours. It reduces the power required to send data per round as the energy draining is spread equally among all nodes. In, PEGASIS energy conservation is achieved in two ways:

1. The head node receives at most two data messages.
2. The distance over which the data are transmitted to closest neighbour is much smaller

So, PEGASIS conserves energy by reducing the number of data messages gathering at head node [8] [9].

g) CCS

CCS is a protocol [16] which reduces energy consumption and extension of PEGASIS protocol. In CCS, the entire network is separated into co-centric circular path and each one of these paths form a cluster. Each path is assigned with a stage. For example, the nearest path to the BS is assigned as stage-1, and as it moves further from the BS the level number increases like Stage-2, stage-3 and so on. In every path, nodes form a chain exactly like PEGASIS. A head node is selected among all of the nodes in the chain and these head nodes are allocated with node numbers. All non head node in a chain, obtain data from its immediate neighbour, aggregate it with its own data and then broadcast it to its immediate neighbour. So it's clear that the head node in each path receives almost two messages. After broadcasting data in a path and receiving it at the head node and then the head nodes in closest path cooperate and send data to the BS. For example the head node in stage-n send out data to the head node in stage-(n-1) and this process persist until sending data to the BS is ended. Data aggregation can be done at every head nodes.

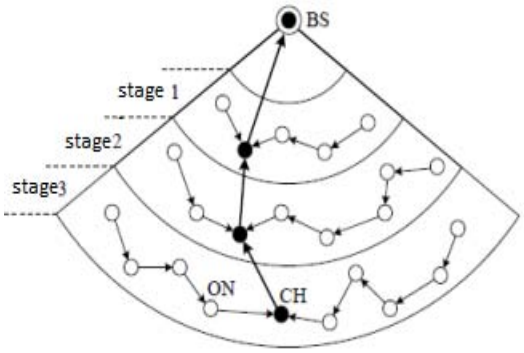


Figure 4 : Stages in CCS algorithm

In this scheme, the distance over source BS from the head node is reduced. This reduced transmission distance saves a significant amount of energy. Also, as the network is separated into a number of concentric clusters, the backward flow of data from BS, which was significant in PEGASIS, is reduced. Due to this, a considerable amount of energy is preserved during data transmission, but this repeated data broadcast can still be less than this protocol proposed.

V. Conclusion

A comparison between various clustering protocol is concluded in the below mention table

Table 1 : comparison of various clustering protocols on various parameters

Protocol	Scalibility	Cluster Stability	Energy Efficiency	Load Balancing	Algorothm Complexity	Delivery Delay
Leach	Very Low	Moderate	Very Low	Moderate	Low	Very Small
Heed	Moderate	High	Moderate	Moderate	Moderate	Moderate
Eecs	Low	High	Low	Moderate	Very High	Small
Pegasis	Very Low	Low	Very High	Moderate	High	Very Large
Teen	Low	Very Low	Moderate	Good	High	Small
Apteen	Low	Low	Low	Moderate	Very High	Small
Ccs	Low	High	Low	Very Bad	Moderate	Large

In this document we have studied the current state of hierarchical routing algorithms, with respect to their various requirements such as energy utilization, stability, delivery delay etc.

In wireless sensor networks, the nodes have limited energy in them which demand to have a careful approach in designing and implementation of clustering algorithm [1]. Moreover there is much future work to be done. Further improvements on energy utilization can be obtained by minimizing the energy used in the clusterhead election process [13]. Energy efficient clustering should eradicate all operating cost associated with the clusterhead selection, as well as with node association respective to their clusterheads. Various algorithms explained the concept of reliability of Sensor network reliability by using re-clustering that occurs in time period; but mostly are energy inefficient and restrict the time accessible within network for data sensing and transmission. Reliability further should be improved by modifying the re-clustering mechanisms subsequent the initial clusterhead selection. Thus reliability can be increase by reducing the wastage and efficient utilization of resources.

References Références Referencias

1. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. Wireless sensor networks: A survey. Comput. Netw. 2002, 38, 393–422.
2. Abbasi, A.A.; Younis, M. A survey on clustering algorithms for wireless sensor networks. Comput. Commun. 2007, 30, 2826–2841.
3. Arboleda, L.M. C.; Nasser, N. Comparison of Clustering Algorithms and Protocols for Wireless Sensor Networks. In Proceedings of IEEE CCECE/CCGEI, Ottawa, ON, Canada, 7–10 May 2006; pp. 1787–1792.
4. Kumarawadu, P.; Dechene, D.J.; Luccini, M.; Sauer, A. Algorithms for Node Clustering in Wireless Sensor Networks: A Survey. In Proceedings of 4th International Conference on Information and Automation for Sustainability, Colombo, Sri Lanka, 12–14 December 2008; pp. 295–300.
5. Deosarkar, B.P.; Yada, N.S.; Yadav, R.P. Cluster Head Selection in Clustering Algorithms for Wireless Sensor Networks: A Survey. In Proceedings of the 2008 International Conference on Computing, Communication and Networking, Virgin Islands, USA, 3–7 August 2008; pp. 1–8.
6. Jiang, C.; Yuan, D.; Zhao, Y. Towards Clustering Algorithms in Wireless Sensor Networks— A Survey. In Proceedings of IEEE Wireless Communications and Networking Conference, Budapest, Hungary, 5–8 April 2009; pp. 1–6.
7. Maimour, M.; Zeghilet, H.; Lepage, F. Cluster-based Routing Protocols for Energy-Efficiency in Wireless Sensor Networks. Available online: http://cdn.intechweb.org/pdfs/12423.pdf (accessed on 14 December 2010).
8. Lotf, J.J.; Hosseinzadeh, M.; Alguliev, R.M. Hierarchical Routing in Wireless Sensor Networks: A Survey. In Proceedings of 2010 2nd International Conference on Computer Engineering and

Technology, Chengdu, China, 16–18 April 2010; pp. 650–654.
9. Boyinbode, O.; Le, H.; Mbogho, A.; Takizawa, M; Poliah, R. A Survey on Clustering Algorithms for Wireless Sensor Networks. In Proceedings of 2010 13th International Conference on Network-Based Information Systems, Takayama, Japan, 14–16 September, 2010; pp. 358–364.
10. Wei, C.; Yang, J.; Gao, Y.; Zhang, Z. Cluster-Based Routing Protocols in Wireless Sensor Networks: A Survey. In Proceedings of 2011 International Conference on Computer Science and Network Technology, Harbin, China, 24–26 December 2011; pp. 1659–1663.
11. Xu, D.; Gao, J. Comparison study to hierarchical routing protocols in wireless sensor networks. Procedia Environ. Sci. 2011, 10, 595–600.
12. Joshi, A.; Lakshmi Priya, M. A Survey of Hierarchical Routing Protocols in Wireless Sensor Network. Available online: http://it.mesce.ac.in/icist/sites/default/files/67-71.pdf (accessed on 17 May 2011).
13. O. Younis and S. Fahmy, "HEED: A Hybrid Energy-Efficient Distributed Clustering Approach for Ad Hoc Sensor Networks," IEEE Transactions on Mobile Computing, vol. 3, no. 4, Oct-Dec 2004.
14. Loscri, G. Morabito, and S. Marano, A Two-Level Hierarchy for Low Energy Adaptive Clustering Hierarchy, DEIS Department, University of Calabria.
15. W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy Efficient Communication Protocol for Wireless Sensor Networks," Proceedings of the33th Hawaii International Conference on System Sciences, 2000.
16. S. M. Jung ,Y. J. Han, and T. M. Chung "The Concentric Clustering Scheme for Efficient Energy consumption in the PEGASUS 9th International Conf. on Advanced Communication Technology, vol.1, pp. 260-265, Feb. 2007
17. A. Manjeshwar and D. P. Aggrawal, TEEN: A Protocol for Enhanced Efficiency in Wireless Sensor Network 1st international Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, 2001, p.189.
18. A. Manjeshwar and D. P. Aggrawal, APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks, 2nd International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile Computing, 2002, 195-202.

This page is intentionally left blank

Global Journal of Computer Science and Technology: E Network, Web & Security
Volume 14 Issue 1 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

On the Investigation of Biological Phenomena through Computational Intelligence

By Jyotsana Pandey & Dr. Bipin Kumar Tripathi
Singhania University, India

Abstract- This paper is largely devoted for building a novel approach which is able to explain biological phenomena like splicing, promoter gene identification, disease and disorder identification, and to acquire and exploit biological data. This paper also presents an overview on the artificial neural network based computational intelligence technique to infer and analyze biological information from wide spectrum of complex problems .Bioinformatics and computational intelligence are new research area which integrates many core subjects such as chemistry, biology, medical science, mathematics, com-puter and information science. Since most of the problems in bioinformatics are inherently hard, ill defined and possesses overlapping boundaries. Neural networks have proved to be effective in solving those problems where conventional com-putation tools failed to provide solution. Our experiments demonstrate the endeavor of biological phenomena as an effec-tive description for many intelligent applications. Having a computational tool to predict genes and other meaningful in-formation is therefore of great value, and can save a lot of expensive and time consuming experiments for biologists. This paper will focus on issues related to design methodology comprising neural network to analyze biological information and investigate them for powerful applications.

Keywords: *splicing, promoter gene, bioinformatics, biological disorder, neural networks.*

GJCST-E Classification : *F.4.1*

Strictly as per the compliance and regulations of:

On the Investigation of Biological Phenomena through Computational Intelligence

Jyotsana Pandey [α] & Dr. Bipin Kumar Tripathi [σ]

Abstract- This paper is largely devoted for building a novel approach which is able to explain biological phenomena like splicing, promoter gene identification, disease and disorder identification, and to acquire and exploit biological data. This paper also presents an overview on the artificial neural network based computational intelligence technique to infer and analyze biological information from wide spectrum of complex problems .Bioinformatics and computational intelligence are new research area which integrates many core subjects such as chemistry, biology, medical science, mathematics, computer and information science. Since most of the problems in bioinformatics are inherently hard, ill defined and possesses overlapping boundaries. Neural networks have proved to be effective in solving those problems where conventional computation tools failed to provide solution. Our experiments demonstrate the endeavor of biological phenomena as an effec-tive description for many intelligent applications. Having a computational tool to predict genes and other meaningful information is therefore of great value, and can save a lot of expensive and time consuming experiments for biologists. This paper will focus on issues related to design methodology comprising neural network to analyze biological information and investigate them for powerful applications.

Keywords: splicing, promoter gene, bioinformatics, biological disorder, neural networks.

I. Introduction

The past few decades have seen a rapid growth in biologi-cal information that is coming in the form of genomes, protein sequences, gene expression, biological disorders data and many other medical diagnosis problems. There is the absolute need of effective and efficient computational tools to store, analyze and interpret the multifaceted data. The conventional techniques [1] of computational biology [2] involve the use of applied mathematics, informatics, statistics and biochemistry to solve biological problems usually on the molecular level. Major research efforts in the field include sequence alignment, gene finding, genome assembly, protein structure alignment, protein structure prediction, and prediction of gene expression, pro-tein-protein interactions and the modeling of evolution. All these problems need to deal with a huge amount of multi-faceted data. For example: there are approximately 26 billion base pairs (bp) representing the various genomes available on the server of the National Center for Biotech-nology Information (NCBI).

The computational biology [2] is concerned with the use of computation to understand biological phenomena and to acquire and exploit biological data, increasingly large-scale data [9]. Methods from computational biology are increa-singly used to augment or leverage traditional laboratory and observation-based biology. These methods have become critical in biology due to recent changes in our ability and determination to acquire massive biological data sets, and due to the ubiquitous, successful biological insights that have come from the exploitation of those data. This transformation from a data-poor to a data-rich field began with DNA sequence data, but is now occurring in many other areas of biology. The bioinformatics involve the crea-tion and advancement of algorithms using techniques in-cluding modern computer science, applied mathematics, statistics, and biochemistry. Hence, in other words, bioin-formatics can be described as the application of computa-tional methods to make biological discoveries [6].

The Computational intelligence [3] is now became a well-established paradigm for solving complex problems dealing with large scale data which are having overlapping, inexact and ill-defined boundaries. Now days, researchers are evolving new theories with a sound biological under-standing in solving problems of molecular and computa-tional biology [9]. They are able to perform a variety of tasks that are difficult or impossible to do with convention-al mathematics, statistics and informatics [12]. To name a few, Tasoulis et al. [10] introduced the application of neural networks, evolutionary algorithms and clustering algo-rithms to DNA microarray experimental data analysis; Liang and Kelemen [11] propose a time lagged recurrent neural network with trajectory learning for identifying and classifying gene functional patterns from the heteroge-neous nonlinear time series microarray experiments.

In this paper we are investigating the method and technique of machine learning through artificial neural networks which proved to be more suitable for genomic and other biological data analysis. The performance of the gene prediction approaches [4] mostly depends on the effectiveness of detecting the splice sites. This paper proposes a system for utilizing an artificial neural network [6] to addresses the problem

Author α: Dept. of Computer Science, Singhania University, Rajasthan, India. e-mail: imjyotsanapandey@gmail.com
Author σ: Dept. of Computer Science & Engineering, HBTI, Kanpur, India. e-mail: abkt.iitk@gmail.com

of splice site detection. ANN takes up it as a two-class problem and classifies a given sequence whether it will be a donor or an acceptor site. Further it predicts the splice form for a given sequence using the scores provided by the single site detectors for every appearing AG and GT dimer. The challenge is to find a splice form that consis-tently combines all predictions. The empirical analysis has further revealed that the results come out more refined if data analyzed in binary format as compared to other format. In the neural network structure, a standard three layer feed forward network of neurons is considered for analysis in which there are two neurons in output corres-ponding to the donor and acceptor splice sites, 128 neurons in hidden layer and 240 units at input end. The 240 input units were used since the orthogonal input scheme uses four inputs each nucleotide in the window.

To provide useful insights for neural network applications in biological information analysis, we structure the rest of the paper as follows: section 2 elaborates the related recent trends in biological information that is coming in the form of genomes [4], protein sequences, gene expression, bio-logical disorders data and medical diagnosis problems. Ar-tificial neural network technique involved in classification and recognition process is presented in section 3. Section 4 presents the empirical evaluation of different biological data analysis and experimental outcome. Finally, section 5 summarizes the paper with the inferences and discussions.

II. Recent Trends in Biological Information

Gene prediction [4] is a very powerful and important task for many ongoing researches in the field of bioinformatics [5]. A gene is a set of instruction which governs the assem-bly and function of all organisms. We know that a gene is a region of DNA that control a certain basic characteristic and ultimately lead to protein synthesis. In the 1990s ge-nomic data started becoming available. Since conventional mathematical models [8] proved to be unworkable in anal-ysis of biological information, bioinformaticians turned to computational intelligence [[9] models for help in tasks such as gene finding and protein structure prediction. The feature selection and class prediction, two learning tasks that are strictly paired in the search of molecular profiles from microarray data, were performed with ANN. The models with ANN have been shown to present a good choice, thus providing analysis and clues for biological information samples. Recently, proteomic data considered potentially rich, but arguably unexploited, for genome an-notation using ANNs which shows favorable performances as compared to conventional mathematical models. The idea of using manifold learning for feature reduction com-bined with an ANN classifier was successful applied in biomedical diagnosis and protein identification.

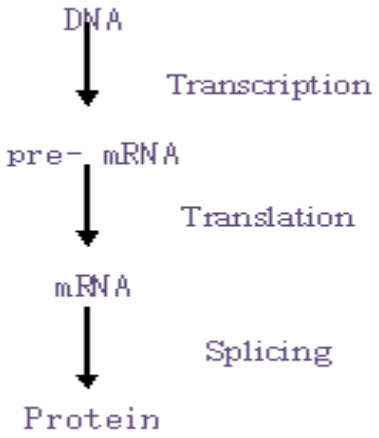


Figure 1 : It shows the importance of splicing which ul-timate cause for making of protein.

In this fig DNA expresses the gene product that it encodes. Figure demonstrates that certain region of the DNA is tran-scribed into RNA in the form of pre-mRNA. Further, the introns of the pre-mRNA are excised, leaving only exon intact to become the mature mRNA by translation. The ribosome then translates the mRNA into a polypeptide chain of amino acids that eventually becomes a protein by splicing [1]. In splice site prediction in E. coli gene DNA se-quences one need to identify the boundaries between exon (the part of DNA sequence retained after splicing) and in-trons (the part of DNA sequence that are spliced out) in given DNA gene sequence. Thus, this problem contains three classes. First is intron–exon (IE) boundary (donors), second is exon-intron (EI) (acceptors) and third class be-longs to neither donors nor acceptors (Neither).

DNA splice sites (Figure 2) are boundaries where splicing occurs and are found between the regions of DNA that code for gene products (exon) and those that do not (intron) [2]. The presence of introns in eukaryotic organisms are believed to be involved in exon shuffling (or alternative splicing) that is responsible for the higher diversity of gene.

Products found in eukaryotic organisms than that of prokaryotic organisms [3]. A typical example of exon shuffling is the generation of antibodies against foreign antigens that may invade the host system. The dinucleotide AG are splice sites that borders the transition from intron to exon (Intron/Exon border) going from 5' to 3', while GT are associated with the transition from exon to intron (Exon/Intron border). The GT dinucleotide is usually referred to as "donor" whereas the AG dinucleotide is known as "acceptor" [4].

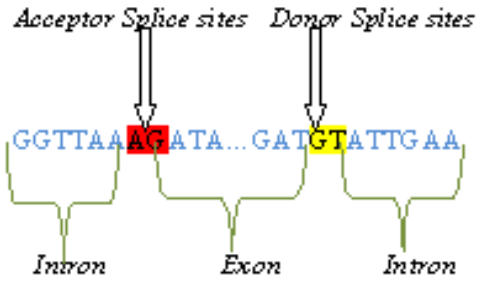


Figure 2 : Schematic representation of the splice site.

III. Neural Network for Biological in-Formation Analysis

Neural networks have several unique characteristics and advantages as tools for the molecular sequence analysis problem. A very important feature of these networks is their adaptive nature, where "learning by example" replaces conventional mathematical techniques which are time-consuming, computation extensive, and weak to noise. A small complexity, robust performance, and quick convergence of artificial neural network (ANN) are vital for its wide applicability. This feature makes such computational models [10] very appealing in application domains where one has little or incomplete understanding of the problem to be solved, but where training data are readily available. Owing to the large number of intercon-nections between their basic processing units, neural networks are error-tolerant, and can deal with noisy data. Neural network [12] architecture encodes information in a distributed fashion. This inherent parallelism makes it easy to optimize the network to deal with a large volume of data and to analyze numerous input parameters. Flexible encoding schemes can be used to combine heterogeneous sequence features for network input. Finally, a multilayer network is capable of capturing and discovering high-order correlations and relationships in input data. The artificial neural networks [13] are "neural" in the sense that they may have been inspired by neuroscience but not necessarily because they are faithful models of biological neural or cognitive phenomena.

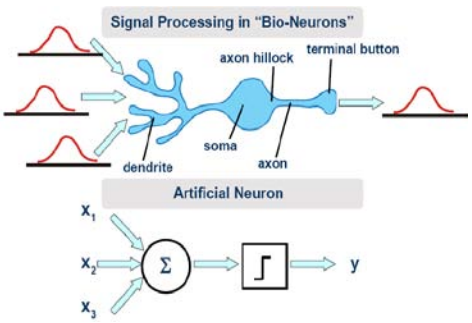


Figure 3 : Real biological neuron to artificial neuron

To enable understanding of neural networks, we start from a theoretical model of a single neuron and then briefly in-troduce a neural network to reveal their structure, training mechanism, operation, and functions. The basic structure of a biological neuron and corresponding artificial neuron is shown in Fig 3 and can be theoretically modeled as (1)

$$Y = f\left(\sum_{i=1}^{n} w_i x_i + b\right), \quad (1)$$

Where X {xi, i=1, 2, . . ., n} represent the inputs to the neu-ron and Y represents the output. Each input is multiplied by its weight wi, a bias b is associated with each neuron and their sum goes through a activation function f. A neural network is characterized by (1) its pattern of connections between the neurons (its architecture), (2) its method of determining the weights on the connections (training or learning, algorithm), and (3) its activation function.

In summary, the applications of ANNs in biological infor-mation processing have to be analyzed individually. ANN has been applied to biological data to deal with the issues that cannot be addressed by traditional algorithms or by other classification techniques. By introducing artificial neural networks, algorithms developed for processing and analysis often become more intelligent than conventional techniques. While neural networks are undoubtedly power-ful tools for classification, clustering and pattern recogni-tion; analysis of the internal weight and bias values for neurons in a network is possible, and a network itself can be represented formulaically, they are sometimes too large to be explained in a way that a human can easily under-stand. Despite this, they are still widely used in situations where a black-box solution is acceptable, and where em-pirical evidence of their accuracy is sufficient for testing and validation.

IV. Design of Learning Machine

It has been widely observed that in comparison to other machine learning approaches [3] neural networks have many positive characteristics for a prospective user. The variety of different network architectures and learning pa-radigms available, coupled with a theoretically limitless number of combinations of layers amounts, connections topologies, transfer functions and neuron amounts, make ANNs incredibly flexible processing tools. They can be applied to data with almost any number of inputs and out-puts, and are well supported in different programming lan-guages and software suites. Through manual modification of weights prior to training, and through imposing custom limitations on their modification during training, existing expert knowledge can be incorporated into their design and construction. Additionally, neural networks based learning machine are usually computationally inexpensive to use after they

have been trained, making them ideal for real-time applications where immediate output is desirable.

The neural networks used in this study (Fig. 3) are of the multi-layer neural network containing neurons of summa-tion aggregation function [13]. They are feed-forward con-nected and have three layers: an input layer, one hidden layer and an output layer. In case of gene prediction prob-lems, the network input is a segment of nucleotides from the nucleotide sequence. The output consists of one unit, giving a real valued output between 0.9 and 0.1. Using a threshold this number is interpreted as a category assign-ment for the nucleotide in the input window. The networks were trained by standard error back propagation learning algorithm on two different tasks: (i) detection of coding nucleotides (versus non-coding nucleotides), and (ii) the prediction of splice sites (defined as the first and last Intron, nucleotide respectively). Thus neural network unarguably possess strong potential for output prediction as can be seen by their widespread use in designing learning machine involving modeling and prediction.

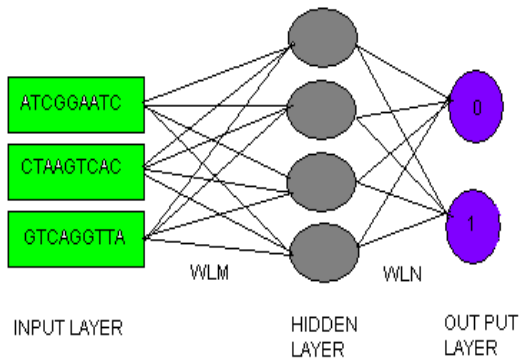


Figure 4 : Learning machine design with artificial neural network

V. Empirical Evaluation of Biological Data Analysis

In order to estimate the strength and effectiveness of bio-logical information from wide spectrum of problems, one needs to analyze them with standard computational intelli-gence technique. I have considered the artificial neural network to prove the motivation and to establish the signi-ficance of work done. Benchmark problems are standard enough to associate the tasks like classification and pattern recognition. They also incorporate the tasks from different fields of importance; few of them are biological engineering, medical and bioinformatics. In this section, I have tho-roughly evaluated kinds of problem to present the impor-tance of neural network in characterizing and analyzing the medical and biological information. In the present investi-gation, we use 5 datasets. Dataset containing primate splice-junction gene sequences and promoter gene se-quences were used in both normalized and binary forms. We observed better error convergence in binary form than the normalized form of dataset. The other datasets Heart Spectf, Bupa Liver Disorder and Protein Localization sites [15] are in numeric forms only, therefore sets normalized in prepro-cessing.

In all the experiments, I have divided whole da-taset into two parts: one is training set and second is test-ing set. Performance is analyzed in terms of parameters which are briefly defined as follows:

Training Accuracy (%)

$$= 100 \times \frac{\text{Number of correct matches}}{\text{Total number of samples in training set}}$$

Testing Accuracy (%)

$$= 100 \times \frac{\text{Number of correct matches}}{\text{Total number of samples in test set}}$$

a) Datasets and Significance

In this paper we have used two genomic datasets and three biological disorders data sets. All these data stets are benchmark and available online for research and academic purposes.

1. E. coli promoter gene sequences (DNA) [11] is acquired to predict the member/non-member of class of sequences with biological promoter activity. The dataset contains non-numeric domain of attributes. The attributes are one of the 'a', 'g','t' and 'c' (a=Adenine, b=Guanine, t=Thymine and c=Cytosine). This dataset have been also used by Harley, C. and Reynolds, R. 1987 in "Analysis of E. Coli Promoter Sequences" Nucleic Acids Research.
2. Primate splice-junctions are the points on DNA sequence at which superfluous DNA is removed during the process of protein creation in higher organisms. The splice-junction gene problem is to identify the boundaries between exons and introns in given DNA gene sequence.
3. Heart SPECTF Data set [14] is based on cardiac single proton emission computed tomography (SPECT) images. Each patient is classified in normal or abnormal categories. Database was used in automated Cardiac SPECT Diagnosis.
4. BUPA liver disorders dataset contains 345 instances that are basically records of 345 males who have taken excessive alcohol consumption. The first 5 attributes are all blood tests which are on thought to be sensitive to liver disorders that might arise from excessive alcohol consumption. The last 2 attributes are different from blood tests. One of them is no of drink having taken in a day and other is selector field.
5. Protein localization sites dataset [15] can be achieve from "Expert System for Predicting Protein

Localization Sites in Gram-Negative Bacteria", Kenta Nakai & Minoru Kanehisa, PROTEINS: Structure, Function, and Genetics 11:95-110, 1991.

b) Learning Machine with Benchmark Datasets

i. *Splice site prediction in E. coli gene DNA sequences*

In E. coli promoter gene sequences (DNA) dataset from University of Wisconsin Biochemistry Department, there are 106 instances with 59 attributes. In these 59 attribute One of {+/-}, indicating the class ("+" = promoter) and second 2-60 remaining 59 fields are the sequence, starting at position filled by one of {a, g, t, c} base pairs. There is no missing attribute Values. In Class attribute there are 53 positive instances and 53 negative instances.

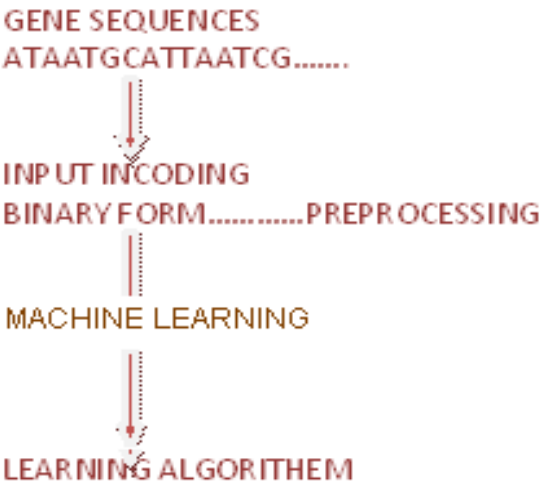


Figure 5 : Processing of biological dataset (binary form) through ANN.

At first step we get DNA sequences to analyze splice sites at second step we used sparse encoding to encode these sequences that is A as (1000), C as (0100), G as (0010) and T as (0001) for preprocessing. It is used to avoid alge-braic dependencies between nucleotides in the encoding also called BIN4 encoding in which each letter coded by four digits with the combination of 0 and 1 to input data. The learning in the neural network is done with error back propagation method and result is presented in Table 1 with discussion in section 6.

ii. *Primate splice-junction gene sequences*

In this dataset all examples taken from Gen bank 64.1 (ftp site: genbank.bio.net), there are three categories "ei", "ie" and "n" for splice sites recognition. Dataset contains 3190 instances including three classes. Class 'EI' as donor con-sists of 767 instances, class 'IE' as accepter consists of 768 instances and rest of as 'N' neither belongs to any class consists of 1655 instances. In this dataset containing pri-mate splice junction gene sequences (DNA), the standard result was 85% and we achieved the accuracy of 81% using binary form of the dataset and 79% of accuracy was achieved when we used the normalized form of dataset. Result is presented in Table 1 with discussion in section 6 for back propagation neural network.

iii. *Heart Spectf Dataset*

This dataset describes diagnosing of cardiac Single Proton Emission Computed Tomography (SPECT) images. This can be achieved from University of Colorado at Denver, Denver, CO 80217, u.s.a.krys.cios@cudenver.edu. Data-base used by Kurgan, L.A., Cios, K.J., Tadeusiewicz, R., Ogiela, M. & Goodenday, L.S. "Knowledge Discovery Ap-proach to Automated Cardiac SPECT Diagnosis" Artificial Intelligence in Medicine, vol. 23:2, pp 149-169, Oct 2001. There are 267 instances as SPECT image sets for patients. Each of the patients is classified into two categories: nor-mal (0) and abnormal (1). Database contains 23 attributes in which 22 spectf image + 1 class. All dataset is divided into training data with 80 instances and testing data with 187 instances. Class 0 consists of 55 instances and class 1 consists of 212 instances. In dataset containing SPECTF heart data the standard result was an accuracy of 87% and we achieve an accuracy of 84% as shown in Table 1.

iv. *BUPA liver disorders data set*

This dataset achieved from "Expert Sytem for Predicting Protein Localization Sites in Gram-Negative Bacteria", Kenta Nakai & Minoru Kanehisa, PROTEINS: Structure, Function, and Genetics 11:95-110, 1991. There are 336 instances each with 8 attributes one of them is name and other are predictive. In this 8 classes are according to the protein location in bacteria. In protein localization dataset containing the standard result was an accuracy of 84% with ad hoc structured probability model; but, we found an ac-curacy of 81% with the artificial neural network.

VI. Inferences and Discussion

Research in bioinformatics is driven by the expe-rimental data. Current biological databases are populated by vast amounts of experimental data. Machine learning has been widely applied to bioinformatics and has gained a lot of success in this research area. At present, with various learning algorithms available in the literature [16], researchers are facing difficulties in choosing the best me-thod that can apply to their data. We performed an empiri-cal study and observed that single learning networks are perfectly usable in splice site prediction, gene prediction, liver disorders and localization site in the same manner. The performance of the learning technique is highly de-pendent on the nature of the training data or on the basis of dataset design. We

conclude that, if dataset is in norma-lized form as well as in binary, then the best results can be achieved.

In the following Table 1, the dataset containing promoter gene sequences (DNA) we achieved the accuracy of 85% using binary form of the dataset and 83% of accu-racy was achieved when we used the normalized form of normal dataset. We can infer that the variation in the results due to the different forms of dataset was because in binary form the variables A-T-G-C are converted into orthogonal vectors. The dataset containing primate splice junction gene sequences (DNA), we achieved the accuracy of 81% using binary form of the dataset and 79% of accuracy was achieved when we used the normalized form of dataset. In the past usage as results of study indicate that machine learning techniques (neural networks, nearest neighbor, contributors' KBANN system) have performed as well/ better than classification based on canonical pattern matching. In dataset containing SPECTF heart data we achieve an accuracy of 84%. In protein localization dataset we found an accuracy of 81% with the artificial neural network. In dataset containing BUPA Liver disorders we found a mediocre accuracy of 77%. We conclude that we achieved the said accuracies with most straight forward and convenient technique which does not possess the complicated computing operations. There may be techniques which may yield little more accuracy for corresponding dataset but our technique is computationally efficient.

We can see the overall analysis at a glance in the Table 1:

Table 1 : Analysis of all datasets with accuracy using back propagation

SNo.	Name of database	No of instances	No of attribute	Training accuracy	Testing accuracy
1	Promoter gene	106	59	85%	83%
2	Primate splice-junction gene sequence	3190	1595	81%	85%
3	SPECTF-heart data	267	23	87%	84%
4	BUPA-liver disorders	345	7	77%	80%
5	Protein Localization Sites.	336	8	81%	84%

VII. Acknowledgements

I want to pay my sincere thanks to all of them who are related directly and indirectly with my work and reviewed & encouraged all the time.

References Références Referencias

1. Krogh A, Brown M, Mian IS, K Sjölander of molecular biology, – Elsevier 1994.
2. Computational systems biology, H Kitano - Nature, 2002.
3. JSR Jang, CT Sun, E Mizutani, -"Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence Automatic Control", IEEE …, ieeexplore.ieee.org 1997.
4. Pati, NN Ivanova, N Mikhailova, G Ovchinnikova Gene-PRIMP:" A gene prediction improvement pipeline for prokaryotic genomes" Nature …, 2010
5. Pierre Baldi, Søren Brunak Bioinformatics: "The Machine Learning Approach", Second Edition (Adaptive Computation and Machine Learning) on Amazon.com. 2002.
6. Zoheir Ezziane, "Applications of artificial intelligence in bioinformatics" 2006.
7. Mitra S and Hayashi Y, "Bioinformatics with soft computing", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol. 36, pp. 616-635 2006..
8. Tasoulis D.K., Plagianakos V. P., and Vrahatis M. N. Com-putational Intelligence Algorithms and DNA Microarrays, Studies in Computational Intelligence (SCI) 94, pp. 1-31. 2008.
9. Tasoulis D.K., Plagianakos V. P., and Vrahatis M. N. Com-putational Intelligence Algorithms and DNA Microarrays, Studies in Computational Intelligence(SCI) 94, pp. 1-31. 2008.
10. Arpad Kelemen Ajith Abraham Yuehui Chen (Eds.) "Com-putational Intelligence in Bioinformatics. Studies in Com-putational Intelligence", Springer-Verlag Berlin Heidelberg, 2008.
11. C. and Reynolds, R. "Analysis of E. Coli Promoter Sequences" Nucleic Acids Research, 15: 2343-2361, 1987.
12. Wu CH, McLarty JW. "Neural Networks and Genome Informatics". Methods in computational Biology and Biochemistry, 1 ed. Vol. 1. Elsevier; 2000.
13. Reese MG, Eeckman FH. "Novel Neural Network Prediction. Systems for Human Promoters And. Splice Sites". 1995.
14. Mariofanna G. Milanova3, Tomasz G. Smolinski4, 20. Ogiela, M. & Goodenday, L.S. "Knowledge Discovery Approach to Automated Cardiac SPECT Diagnosis" Artificial Intelligence in Medicine, vol. 23:2, pp 149-169, Oct 2001.
15. Kenta Nakai & Minoru Kanehisa, "Expert Sytem for Predicting Protein Localization Sites in Gram-

Negative Bacteria", PROTEINS: Structure, Function, and Genetics 11: 95-110, 1991.

16. Aik Choon TAN and David GILBERT, “An empirical com-parison of supervised machine learning techniques in bioin-formatics” Bioinformatics Research Centre, Department of Computing Science 12 Lilybank Gardens, University of Glasgow, Glasgow G12 8QQ, UK 2009.

Global Journals Inc. (US) Guidelines Handbook 2014

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

Global Journals Incorporate (USA) is accredited by Open Association of Research Society (OARS), U.S.A and in turn, awards "FARSC" title to individuals. The 'FARSC' title is accorded to a selected professional after the approval of the Editor-in-Chief/Editorial Board Members/Dean.

The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

FARSC accrediting is an honor. It authenticates your research activities. After recognition as FARSC, you can add 'FARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, and Visiting Card etc.

The following benefits can be availed by you only for next three years from the date of certification:

FARSC designated members are entitled to avail a 40% discount while publishing their research papers (of a single author) with Global Journals Incorporation (USA), if the same is accepted by Editorial Board/Peer Reviewers. If you are a main author or co-author in case of multiple authors, you will be entitled to avail discount of 10%.

Once FARSC title is accorded, the Fellow is authorized to organize a symposium/seminar/conference on behalf of Global Journal Incorporation (USA).The Fellow can also participate in conference/seminar/symposium organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent.

You may join as member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. In addition, it is also desirable that you should organize seminar/symposium/conference at least once.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The FARSC can go through standards of OARS. You can also play vital role if you have any suggestions so that proper amendment can take place to improve the same for the benefit of entire research community.

As FARSC, you will be given a renowned, secure and free professional email address with 100 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.

The FARSC will be eligible for a free application of standardization of their researches. Standardization of research will be subject to acceptability within stipulated norms as the next step after publishing in a journal. We shall depute a team of specialized research professionals who will render their services for elevating your researches to next higher level, which is worldwide open standardization.

The FARSC member can apply for grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A. Once you are designated as FARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria. After certification of all your credentials by OARS, they will be published on your Fellow Profile link on website https://associationofresearch.org which will be helpful to upgrade the dignity.

The FARSC members can avail the benefits of free research podcasting in Global Research Radio with their research documents. After publishing the work, (including published elsewhere worldwide with proper authorization) you can upload your research paper with your recorded voice or you can utilize chargeable services of our professional RJs to record your paper in their voice on request.

The FARSC member also entitled to get the benefits of free research podcasting of their research documents through video clips. We can also streamline your conference videos and display your slides/ online slides and online research video clips at reasonable charges, on request.

The FARSC is eligible to earn from sales proceeds of his/her researches/reference/review Books or literature, while publishing with Global Journals. The FARSC can decide whether he/she would like to publish his/her research in a closed manner. In this case, whenever readers purchase that individual research paper for reading, maximum 60% of its profit earned as royalty by Global Journals, will be credited to his/her bank account. The entire entitled amount will be credited to his/her bank account exceeding limit of minimum fixed balance. There is no minimum time limit for collection. The FARSC member can decide its price and we can help in making the right decision.

The FARSC member is eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get remuneration of 15% of author fees, taken from the author of a respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

The ' MARSC ' title is accorded to a selected professional after the approval of the Editor-in-Chief / Editorial Board Members/Dean.
The “MARSC” is a dignified ornament which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., MARSC or William Walldroff, M.S., MARSC.

MARSC accrediting is an honor. It authenticates your research activities. After becoming MARSC, you can add 'MARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, Visiting Card and Name Plate etc.

The following benefitscan be availed by you only for next three years from the date of certification.

MARSC designated members are entitled to avail a 25% discount while publishing their research papers (of a single author) in Global Journals Inc., if the same is accepted by our Editorial Board and Peer Reviewers. If you are a main author or co-author of a group of authors, you will get discount of 10%.

As MARSC, you will be given a renowned, secure and free professional email address with 30 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The MARSC member can apply for approval, grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A.

Once you are designated as MARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria.

It is mandatory to read all terms and conditions carefully.

AUXILIARY MEMBERSHIPS

Institutional Fellow of Open Association of Research Society (USA)-OARS (USA)

Global Journals Incorporation (USA) is accredited by Open Association of Research Society, U.S.A (OARS) and in turn, affiliates research institutions as "Institutional Fellow of Open Association of Research Society" (IFOARS).

The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

The IFOARS institution is entitled to form a Board comprised of one Chairperson and three to five board members preferably from different streams. The Board will be recognized as "Institutional Board of Open Association of Research Society"-(IBOARS).

The Institute will be entitled to following benefits:

The IBOARS can initially review research papers of their institute and recommend them to publish with respective journal of Global Journals. It can also review the papers of other institutions after obtaining our consent. The second review will be done by peer reviewer of Global Journals Incorporation (USA) The Board is at liberty to appoint a peer reviewer with the approval of chairperson after consulting us.

The author fees of such paper may be waived off up to 40%.

The Global Journals Incorporation (USA) at its discretion can also refer double blind peer reviewed paper at their end to the board for the verification and to get recommendation for final stage of acceptance of publication.

The IBOARS can organize symposium/seminar/conference in their country on behalf of Global Journals Incorporation (USA)-OARS (USA). The terms and conditions can be discussed separately.

The Board can also play vital role by exploring and giving valuable suggestions regarding the Standards of "Open Association of Research Society, U.S.A (OARS)" so that proper amendment can take place for the benefit of entire research community. We shall provide details of particular standard only on receipt of request from the Board.

The board members can also join us as Individual Fellow with 40% discount on total fees applicable to Individual Fellow. They will be entitled to avail all the benefits as declared. Please visit Individual Fellow-sub menu of GlobalJournals.org to have more relevant details.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

After nomination of your institution as "Institutional Fellow" and constantly functioning successfully for one year, we can consider giving recognition to your institute to function as Regional/Zonal office on our behalf.
The board can also take up the additional allied activities for betterment after our consultation.

The following entitlements are applicable to individual Fellows:

Open Association of Research Society, U.S.A (OARS) By-laws states that an individual Fellow may use the designations as applicable, or the corresponding initials. The Credentials of individual Fellow and Associate designations signify that the individual has gained knowledge of the fundamental concepts. One is magnanimous and proficient in an expertise course covering the professional code of conduct, and follows recognized standards of practice.

Open Association of Research Society (US)/ Global Journals Incorporation (USA), as described in Corporate Statements, are educational, research publishing and professional membership organizations. Achieving our individual Fellow or Associate status is based mainly on meeting stated educational research requirements.

Disbursement of 40% Royalty earned through Global Journals : Researcher = 50%, Peer Reviewer = 37.50%, Institution = 12.50% E.g. Out of 40%, the 20% benefit should be passed on to researcher, 15 % benefit towards remuneration should be given to a reviewer and remaining 5% is to be retained by the institution.

We shall provide print version of 12 issues of any three journals [as per your requirement] out of our 38 journals worth $ 2376 USD.

Other:

The individual Fellow and Associate designations accredited by Open Association of Research Society (US) credentials signify guarantees following achievements:

- The professional accredited with Fellow honor, is entitled to various benefits viz. name, fame, honor, regular flow of income, secured bright future, social status etc.

- In addition to above, if one is single author, then entitled to 40% discount on publishing research paper and can get 10%discount if one is co-author or main author among group of authors.
- The Fellow can organize symposium/seminar/conference on behalf of Global Journals Incorporation (USA) and he/she can also attend the same organized by other institutes on behalf of Global Journals.
- The Fellow can become member of Editorial Board Member after completing 3yrs.
- The Fellow can earn 60% of sales proceeds from the sale of reference/review books/literature/publishing of research paper.
- Fellow can also join as paid peer reviewer and earn 15% remuneration of author charges and can also get an opportunity to join as member of the Editorial Board of Global Journals Incorporation (USA)
- • This individual has learned the basic methods of applying those concepts and techniques to common challenging situations. This individual has further demonstrated an in–depth understanding of the application of suitable techniques to a particular area of research practice.

Note :

"

- In future, if the board feels the necessity to change any board member, the same can be done with the consent of the chairperson along with anyone board member without our approval.
- In case, the chairperson needs to be replaced then consent of 2/3rd board members are required and they are also required to jointly pass the resolution copy of which should be sent to us. In such case, it will be compulsory to obtain our approval before replacement.
- In case of "Difference of Opinion [if any]" among the Board members, our decision will be final and binding to everyone.

"

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.Online Submission: There are three ways to submit your paper:

(A) (I) First, register yourself using top right corner of Home page then Login. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer, then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Swis 721 Lt BT.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be three lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.
Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads:

Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications.

Research letters: The letters are small and concise comments on previously published matters.

5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also.Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:

Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words

A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

You must strictly follow above Author Guidelines before submitting your paper or else we will not at all be responsible for any corrections in future in any of the way.

Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

1. Choosing the topic: In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. Pick a good study spot: To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. Think technically: Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page
- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.

- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.
- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Administration Rules Listed Before Submitting Your Research Paper to Global Journals Inc. (US)

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.

- Do not give permission to anyone else to "PROOFREAD" your manuscript.

- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

Criterion for Grading a Research Paper (Compilation) by Global Journals Inc. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

Index

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

ISSN 9754350