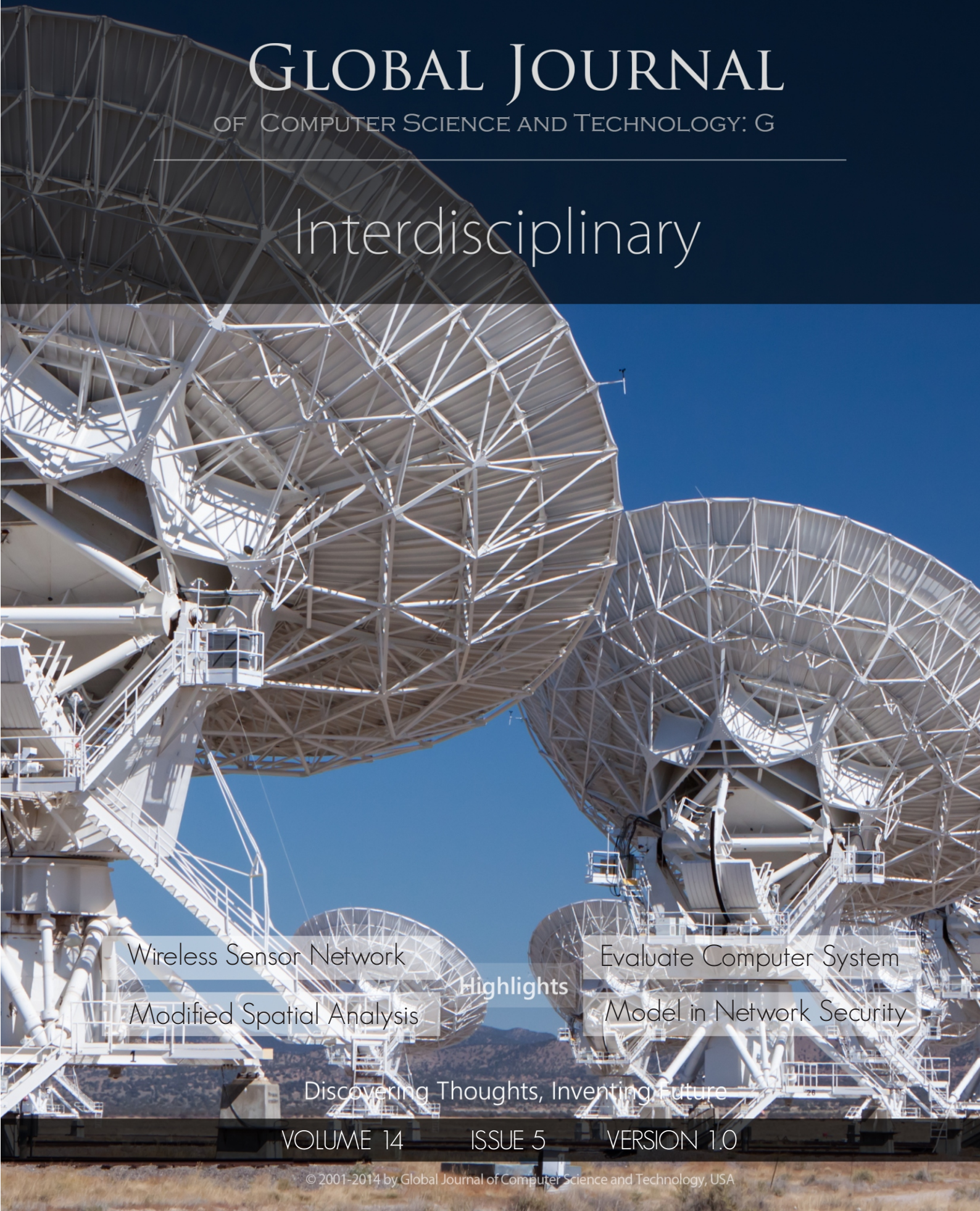
Online ISSN : 0975-4172
Print ISSN : 0975-4350

Global Journal

of Computer Science and Technology: G

Interdisciplinary



Highlights

Wireless Sensor Network

Modified Spatial Analysis

Evaluate Computer System

Model in Network Security

Discovering Thoughts, Inventing Future

VOLUME 14 ISSUE 5 VERSION 1.0

Global Journal of Computer Science and Technology: G
Interdisciplinary

Global Journal of Computer Science and Technology: G
Interdisciplinary

Volume 14 Issue 5 (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 Issue

i. Copyright Notice
ii. Editorial Board Members
iii. Chief Author and Dean
iv. Contents of the Issue

1. A Novel Method of Violated Constraint Prediction with Modified Spatial Analysis based Fuzzy Sorting. *1-5*
2. Using Latency to Evaluate Computer System Performance. *7-15*
3. Analysis of Handoff Latency in Advanced Wireless Networks. *17-22*
4. A Text Mining-based Anomaly Detection Model in Network Security. *23-31*
5. Priority based Congestion Control Mechanism in Multipath Wireless Sensor Network. *33-38*
6. Voip End-to-End Security using S/Mime and a Security Toolbox. *39-42*

v. Fellows and Auxiliary Memberships
vi. Process of Submission of Research Paper
vii. Preferred Author Guidelines
viii. Index

Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 14 Issue 5 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 Novel Method of Violated Constraint Prediction with Modified Spatial Analysis based Fuzzy Sorting

By K. Nithiya & A. Vinoth Kannan
IFET College of Engineering, India

Abstract- Mobility Prediction of a Moving Node and Network Delay is an important performance characteristic of a wireless network. The Data delivery Delay of a network specifies how long it takes for a data to travel across the network from one node or endpoint to another. It is typically measured in multiples or fractions of seconds. The work presented here belongs to domain of data mining cum wireless network , the Real Time Early Prediction of network delay based on mobility is done using the proposed spatial analysis for constraint violation prediction method. A New application is presented concerning the Delivery delays of UDP packets in GPRS network. The GPS points that are collected from GPS module is analyzed using proposed spatial analysis, for future location prediction using Timestamps as primary data .

Index-terms: *monitoring system, delay analysis, GPRS, GPS, UDP/IP, time constraint, map matching.*

GJCST-G Classification: *I.5.1*

Strictly as per the compliance and regulations of:

A Novel Method of Violated Constraint Prediction with Modified Spatial Analysis based Fuzzy Sorting

K. Nithiya[α] & A. Vinoth Kannan[σ]

Abstract- **Mobility Prediction of a Moving Node and Network Delay is an important performance characteristic of a wireless network. The Data delivery Delay of a network specifies how long it takes for a data to travel across the network from one node or endpoint to another. It is typically measured in multiples or fractions of seconds. The work presented here belongs to domain of data mining cum wireless network , the Real Time Early Prediction of network delay based on mobility is done using the proposed spatial analysis for constraint violation prediction method. A New application is presented concerning the Delivery delays of UDP packets in GPRS network. The GPS points that are collected from GPS module is analyzed using proposed spatial analysis, for future location prediction using Timestamps as primary data .**

Index-terms: monitoring system, delay analysis, GPRS, GPS, UDP/IP, time constraint, map matching.

I. Introduction

Real Time Applications usually impose strict time constraints which affect the Grade of service. Time constraints Restricts the Time gap between 2 locations. IP network delays can range from just a few milliseconds to several hundred milliseconds. When two devices communicate with each other using a packet switched network(GPRS), it takes a certain amount of time for information to transmit and receive the data. The total time that it takes for this chunk of information, commonly called a packet, to travel end-to-end is called network delay.

In this Proposed Violation Prediction method, Communication or ope rating delays between 2 datas are bounded and are taken into account by verifying a global time constraint. The uncertainty induced by these delays generates an uncertainty on the verification that results in a possibility measure associated with constraint verification. Freschet Distance (1999) based prediction lack of True path of Moving Object[1]. The performance of Kalman filter approach depends only on the quality of electronic map data and error sources (2011) associated with positioning devices were not considered[13]. Coorelation ananlysis (2012) shift the received signal by delay and multiply it with other series[12]. Even Fuzzy Logic based matching does not consider error sources when estimating the location. Dynamic time windows (2011) based delay estimation based on Kalman Filter restricted to stastitical data [6].

Our Objective is therefore to study the particular problem that whenever vehicle location request is made its current position will not be retrieved accurately, instead its previous position will not be retrieved accurately, instead its previous position alone sent to requested client. For that, we suppose that communication delays between devices are bounded. This uncertainty on communication delays induces an uncertainty on the time constraint verification. The exploitation of the obtained results allows recognising in a distributed way, the occurrence of the failure symptom with a certain possibility. If a target node moves linearly, through zone prediction method we can predict the location accurately. However, on the other side, when a target does not move linearly means changes it"s direction such as though spiral way. When a device on a packet switching network sends information to another device, it takes a certain amount of time for that information, or data, to travel across the network and be received at the other end.This delay still becomes worst when using Unreliable UDP packets.

II. Analysis on Spatial Constraints

Existing localization techniques which mostly rely on GPS technology are not able to provide reliable positioning accuracy in all situations. This spatial based map matching technique will satisfy the real time constraints to reduce data delivery time and further provide accuracy than existing methods. The Important terms used are described in this section.

a. *A GPS log* is a collection of GPS points $L = \{p1, p2, \ldots, pn\}$. Each GPS point $pi \in L$ contains latitude . lat, longitude . lng and timestamp pi . t.

b. *GPS Trajectory:* A GPS Trajectory T is a sequence of GPS points with the time interval between any consecutive GPS points not exceeding a certain threshold ΔT , i.e. T: $p1 \rightarrow p2 \rightarrow \cdots \rightarrow pn$, where $pi \in L$, and $0 < pi{+}1.\ t - pi\ .\ t < \Delta T$ $(1 \leq i < n)$. Figure 3 shows an example of GPS trajectory. ΔT is the sampling interval. In this paper, we focus on low sampling rate GPS trajectories with $\Delta T \geq 5min$ and maximum Back off attempt is 15 times

Author α σ : M.E Student, Assistant Professor ECE Dept Applied Electronics , IFET college of Engineering Villupuram, Tamilnadu.
e-mail : Kn516894@gmail.com.

c. *Road Segment:* A road segment e is a directed edge that is associated with an id $e.eid$, a typical travel speed $e.v$, a length value $e.l$, a starting point $e.start$, an ending point $e.end$ and a list of intermediate points that describes the road using a polyline.

d. *Network:* A network is a directed graph (V, E), where V is a set of vertices representing the intersections and terminal points of the road segments, and E is a set of edges representing road segments.

e. *Path:* Given two vertices , Vj in a road network G, a path P is a set of connected road segments that start at Vi and end at Vj

f. *Timestamps :* Start time of each iterations represent the Timestamp of that data set. This elapsed time for each trajectory is obtained from Tic function using MATLAB.

III. Hardware and Software Requirements

GPS and GPRS module BU353 USB receiver and SIM 300which is a Trib and GSM/GPRS engine works on frequencies EGSM 900\ MHz, DCS 1800 MHz and PCS 1900 MHz is preferred for Getting the Location data such as Latitude Longitude and Timestamps are Analyzed in MATLAB Tool and command using lat_calc(i)=str2double(lat(i,:)) function for 10 iterations. Obtained values are used as references for improving Map matching accuracy. Tic and toc commands are used for internal stopwatch timer interval recordings for each trajectory set.

IV. Proposed Architecture

In this spatial based location prediction method , the target location at an instance of time is predicted based on previous good locations using an iterative process. Once the zone of the target is predicted with respect a relative origin, the Previous location Points from trajectory data is used to find future location and packet delivery speed enhancement . This is done using Timestamps from which the data is sent and time of its arrival. The values are Recorded for each transmission of a packet to find out the transmission time which is the difference between Arrival and sent. WGS 84 coordinate datum is converted to Decimal Degrees as first step to enhance accuracy in prediction.

Consequential Movements is given as Tj, Tj+1□

TJ+1 - Tj ≤ Maximum Time gap.

Proposed Spatio Temporal analysis System

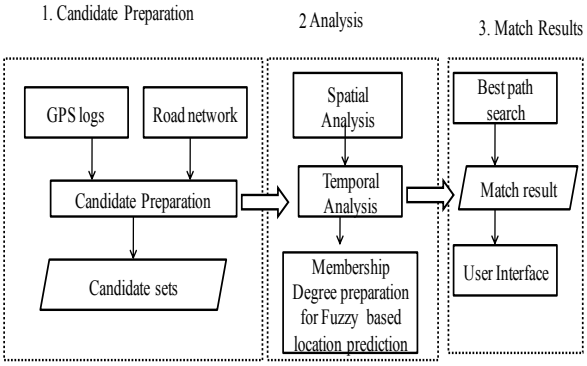


Figure 1: Proposed Spatial Analysis Based Constraint Prediction

The strict Timing constraints has to be satisfied for achieving best map matching accuracy and packet delay analysis. To fulfill this need in networks in which the topology changes frequently, these changes should not affect the Quality of Service (QoS) for data delivery. The maximum gap defined by Floating point Time difference between 2 successive GPS points.

The IP configuration for sending and receiving packets is carried out using UDP socket creation in DOS prompt, since this prediction is carried over Client , Server Architecture.

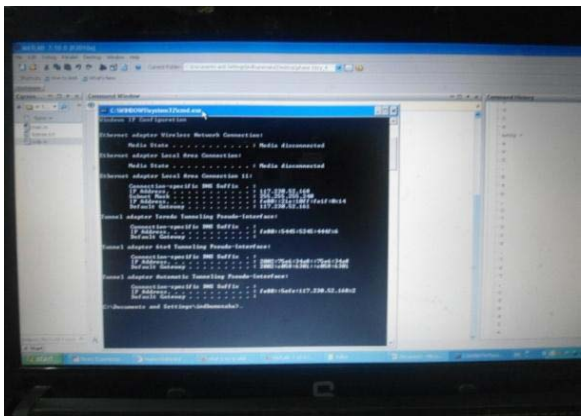


Figure 1(b) : IP configuration using IP Address

Fuzzy Sorting- It is not always possible to do computations with real values., due to unknown GPS noise and error values obtained, if interval between the values related to uncertained sequences are considered for fuzzy sorting. Rules related to satisfy those time Time constraints are created manually and checked for all sequent points

If „n" number of intervals is of form[ai,bi], permutation is created Cj contains all [aij,bij] to satisfy all constraints c1 ≤ c2 ≤ c3 ≤ cn. If the Timestamp of last point when compared with current GPS point is smaller than timestamp of last data of previous trajectory , then Each rule assigned a measure degree α and β to predict the error source using Rectangular Grid over tracking region. Error point over each edge of grid is considered for future mobility prediction.

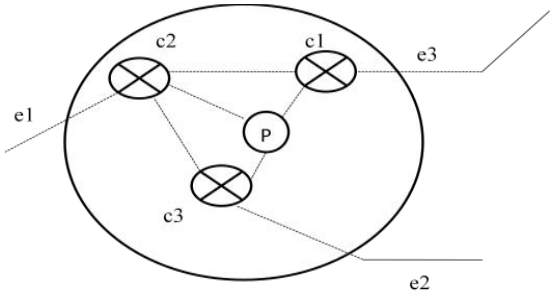


Figure 2 : Trajectory Of Gps Assisted vehicle path

Constraints are expressed by timing relationships between Trajectory(GPS points) A constraint can, for example, express a transport time window between two locations. To determine the Time window constraints consider the observed sample and characteristics of the correctly operating system are used to create a confidence space of possible timing relationships between Gps Trajectory obtained from the system. To execute simulation for zone finding, MATLAB is used as a simulation tool. The UDP packet socket is created using Send and Receive Arguments with IP configuration.

From the given point P, Within radius „ r" canditate projection is a line drawn from point P to the Road side of the segment. Line segment is projected from the point P to the road segment e, and named as C. shortest distance between p and c is the road that vehicle is choosen to travel as an assumption for this the proposed concept. In spatial analysis, both geometric and topological information of the road network is used to evaluate the candidate points are used for Time instance evaluation. If any value greater than max Gap is obtained ,

The observation probability(geometric information) is defined as the likelihood that a GPS sampling point pi matches a candidate point ci computed based on the distance between the two points Transmission Probability(topological information) is used to identify the true path if a cross path is located wrongly. In this if a cross path is identified wrongly, then the previous P point is compared and the path is follower regarding it. The nodal delays accumulate and give an end-to-end delay,

Time difference = Timestamp of packet Received – Packet Sent Time

End- End delay = Arrival time – Received time / number of iterations used. Here 10

The Total Number of iterations used is 10. One of the major concerns of this research is to keep track the moving target as well as stationary target. As the target can move any direction dynamic references used for applying proper geometry in triangulation based map matching method instead of stationary references. That"s"why modified spatial map matching is done here to reduce the execution time of proposed method .The iteration process earned the execution time of 0.0019 sec with respect to

V. Proposed Algorithm based on Time Constraints

The Algorithm defines the each Point"s timing relationship (Timestamps) and their Sequencing relationship.

STEP 1 : Initialize list of candidate points an an empty list

STEP 2 : For i= 1 to n do

STEP 3 : Get candidate values for observed GPS node positions. From the given point P, With in radius r canditarete projection is a line drawn from point P to the road side of the segment. Line segment is projected from the point P to the road segment e, and named as C.

STEP 4 : Time Difference between successive GPS points are recorded in Mantissa format to include temporally similarity with respect to current point received. (Here VB event driven programming language is used)

STEP 5 : Line segment is projected from the point P to the road segment e, and named as C. shortest distance between p and c is the road that vehicle is choosen to travel

STEP 6 : The observation probability (geometric information) is defined as the likelihood that a GPS sampling point pi matches a candidate point ci computed based on the Time difference between the two points

STEP 7 : Transmission Probability(topological information) is used to identify the true path if a cross path is located wrongly. In this if a cross path is identified wrongly, then the previous P point is compared and the path is follower regarding it.

STEP 8 : Binary exponential backoff (truncated exponential backoff) is used to transmit datas data with number of attempts restricted to 15.

STEP 9 : The Candidate graph is constructed to find the relative sequence to be matched for prediction Violations of Constraints due to Measurement errors.

STEP 10 : Return the matched sequence with less delay using membership functions based on

Fuzzy sorting- It takes for last Iteration of data received with 2.400 sec delay and total execution time is 0.0019 sec which is less then the existing map match method Hidden markov model and kalman filter implementation. AT commands such as CIPSTART, CIPCLOSE are used to Initiate and stop GPS device and Time Difference between points obtained from OSM or Google Map can be used to visually represent the violated constraint.

Values of Final Iteration

Sample no	Latitude	Longitude	Delay(sec)	Process Time
1	2400.00000	12100.00000	2.1000 sec	0.00037
2	2400.01000	12100.00000	0.7000 sec	0.00028
3	2400.02000	12100.00000	1.3000 sec	0.00020
4	2400.03000	12100.01000	0.4000 sec	0.00019
5	2400.04000	12100.00000	4.7000 sec	0.00019
6	2400.05000	12100.00000	3.9000 sec	0.00019
7	2400.06000	12100.02000	1.4000 sec	0.00019
8	2400.07000	12100.00000	4.3000 sec	0.00019
9	2400.08000	12100.00000	2.4000 sec	0.00019

Figure 3 : Delay Obtained for packet through Iterated Time constraint analysis

Thus, Delivery delay is determined using Absolute departure time of 1st data and last data by monitoring the time windows. This can be further expanded By Fuzzy based sorting using Membership functions in near future. samples are analysed with MATLAB for defined spacing relationship of GPS points with various sampling rates. If one constraint is found to be violated therefore delay occurs on forthcoming Data deliveries. Hence all the forthcoming coming constraints are to checked on the assumed route with constant sampling rate trajectories.

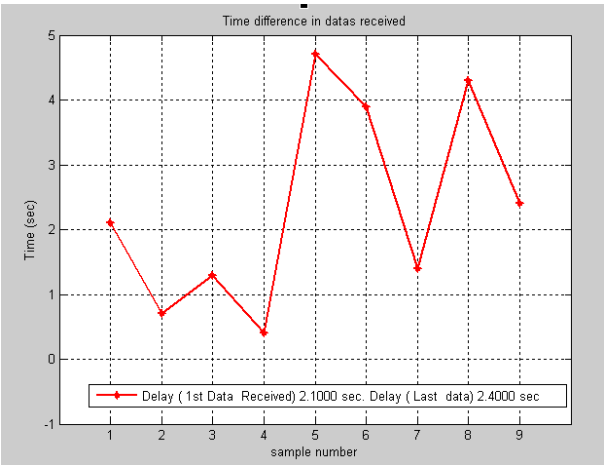


Figure 4 : Time Difference execution using Matlab

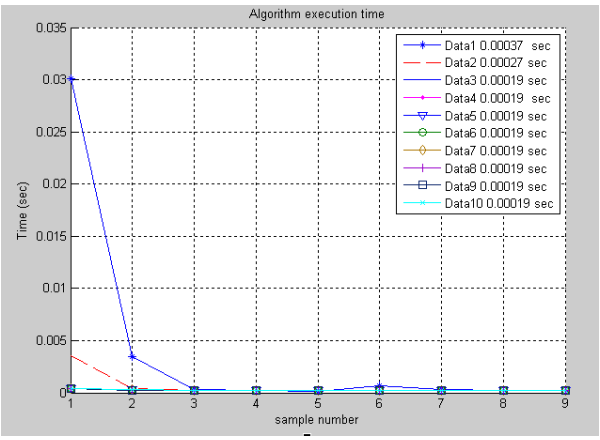


Figure 5 : Observed Sample vs Execution Time for delay prediction

Map matching is a Error correction technique done for pattern identification of latitude, Longitude points and also for packet delivery delay constraints formulation. This method can be useful for any kind of Physical network in Future mainly for Distributed systems, Peer-peer Networks to determine the END-END delay before packet reaches destination in network.

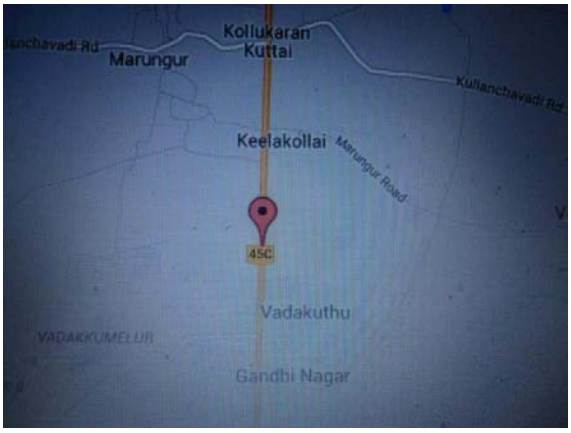


Figure 6 : Time constaint violation monitoring in Google Map near Villupuram, Tamilnadu

VI. Conclusion

This Paper clearly examine the performance of algorithm and datas from correctly observed operating system to predict the delivery delay before the packet Reaches the destination. The time from when the packet is sent and time it received at the Receiver is recorded and it becomes crucial fator more than the threshold value here assumed Maximum Gap of 0.00019 seconds, the time difference is further applied using Fuzzy sorting to observe Error sources. So far, Spatial Analysis Results that are found is presented in this paper and membership function Estimation for Fuzzy based Interval value sorting process is in progress and presented in Future. Delays are predicted earlier within the reach of it to destination using Advanced Time Constraint violation Prediction Algorithm , at the same time by plotting the obtained values in Google map using appropriate interface, up-to date data receival with less delay and more accuracy is achieved by the proposed method. The Future Scope implies that Delay prediction is possible for GPRS, 3G Networks and delay diagnosis can be done by expanding the concept using FUZZY Sorting by proposed Algorithm by 2016.

References Références Referencias

1. Jensen C. S. Capturing the uncertainty of moving-object representations based on Freschet Distance, in Proceedings of the 6th international Symposium on Advances in Spatial Databases, 111-132, 1999
2. Samet, H., Sankaranarayanan, J., and Alborzi, H. Scalable network distance browsing in spatial databases. In Proceedings of the 2008 ACM SIGMOD international Conference on Management of Data, 43-54, 2008
3. Zheng, Y., Chen, Y., Xie, X., and Ma, W. GeoLife2.0: a location-based social networking service. In proceedings of International Conference on Mobile Data Management, 2009.
4. H. Jula, M. Dessouky, and P. A. Ioannou, "Truck route planning in nonstationary stochastic networks with time windows at customer locations," *IEEE Trans. Intell. Transp. Syst.,* vol. 7, no. 1, pp. 51–62, Mar. 2006.
5. S. Murugananham and P. R. Mukesh, "Real time web based vehicle tracking using GPS," *J. World Acad. Sci. Eng. Technol.,* vol. 61, pp. 91– 99,Jan. 2010.
6. W.-H. Lin and D. Tong, "Vehicle reidentification with dynamic time windows for vehicle passage time estimation," *IEEE Trans. Intell. Transp. Syst., vol.* 12, no. 4, pp. 1057–1063, Dec. 2011...
7. R. Raymond, T. Morimura, T. Osogami and N. Hirosue, "Map Matching Hidden Markov Model on Sampled Road Network" In Proceedings of ICPR'12, 2012
8. "Intelligent Data Mining Technique for Intrusion Detection Models On Network" by GM Nazer In European Journal of Scientific Research 71, 2012.
9. E. L. Martelot and C. Hankin. Fast Multi- Scale Detection of Relevant Communities in Large Scale Networks. The Computer Journal, Oxford University Press, 2013
10. Sharmistha Khan, Dr. Dhadesugoor R. Vaman, Siew T. Koay proposed "Sequential prediction of Zone and PL&Tof nodes in mobile networks "in International Journal of Wireless & Mobile Networks (IJWMN) Vol. 6, No. 4, August 2014.
11. Ibrahim Sayed Ahmad , Ali Kalakech , Seifedine Kadry " Modified Binary Exponential Backoff Algorithm to Minimize Mobiles Communication Time"in *I.J. Information Technology and Computer Science,* 2014, 03, 20-29
12. Mohammad Al-Hattab, Maen Takruri and Johnson Agbinya proposed "Mobility prediction based on coorelation analysis using History of log data "International Journal of Electrical & Computer Sciences IJECS-IJENS Vol: 12 No: 03 in 2012
13. Davidson P.J. Collin "Application of Particle filters for map matching" journal of navigations vol2(4),pp.285-292,-2011 Author Profile:

This page is intentionally left blank

Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 14 Issue 5 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

Using Latency to Evaluate Computer System Performance

By Olawuyi J. O., Fagbohunmi S. G., Olawuyi O. M., Mgbole F.
Abia State Polytechnic, Nigeria

Abstract- Building high performance computer systems requires an understanding of the behaviour of systems and what makes them fast or slow. In addition to our file system performance analysis, we have a number of projects in measuring, evaluating, and understanding system performances. The conventional methodology for system performance measurement, which relies primarily on throughput-sensitive benchmarks and throughput metrics, has major limitations when analyzing the behaviour and performance of interactive workloads. The increasingly interactive character of personal computing demands new ways of measuring and analyzing system performance. In this paper, we present a combination of measurement techniques and benchmark methodologies that address these problems. We use some simple methods for making direct and precise measurements of event handling latency in the context of a realistic interactive application. We analyze how results from such measurements can be used to understand the detailed behaviour of latency-critical events. We demonstrate our techniques in an analysis of the performance of two releases of Windows 9x and Windows XP Professional. Our experience indicates that latency can be measured for a class of interactive workloads, providing a substantial improvement in the accuracy and detail of performance information over measurements based strictly on throughput.

GJCST-G Classification: B.8.2

USINGLATENCYTOEVALUATECOMPUTERSYSTEMPERFORMANCE

Strictly as per the compliance and regulations of:

Using Latency to Evaluate Computer System Performance

Olawuyi J.O. [α], Fagbohunmi S.G. [σ], Olawuyi O.M. [ρ] & Mgbole F. [ω]

Abstract- Building high performance computer systems requires an understanding of the behaviour of systems and what makes them fast or slow. In addition to our file system performance analysis, we have a number of projects in measuring, evaluating, and understanding system performances. The conventional methodology for system performance measurement, which relies primarily on throughput-sensitive benchmarks and throughput metrics, has major limitations when analyzing the behaviour and performance of interactive workloads. The increasingly interactive character of personal computing demands new ways of measuring and analyzing system performance. In this paper, we present a combination of measurement techniques and benchmark methodologies that address these problems. We use some simple methods for making direct and precise measurements of event handling latency in the context of a realistic interactive application. We analyze how results from such measurements can be used to understand the detailed behaviour of latency-critical events. We demonstrate our techniques in an analysis of the performance of two releases of Windows 9x and Windows XP Professional. Our experience indicates that latency can be measured for a class of interactive workloads, providing a substantial improvement in the accuracy and detail of performance information over measurements based strictly on throughput.

I. Introduction

Benchmarks are used in computer systems research to analyze design alternatives, identify performance problems, and motivate improvements in system design. Equally important, consumers use benchmarks to evaluate and compare computer systems. Current benchmarks typically report throughput, bandwidth, or end-to-end latency metrics. Though often successful in rating the throughput of transaction processing systems and/or the performance of a system for scientific computation, these benchmarks do not give a direct indication of performance that is relevant for interactive applications such as those that dominate modern desktop computing. The most important performance criterion for interactive applications is responsiveness, which determines the performance perceived by the user.

Author α ω: Department of Computer Science, Abia State Polytechnic, Aba. e-mails: olawuyijo@yahoo.co.uk, mojirayool use ye2014@yahoo.com
Author σ: Department of Computer Engineering, Abia State Polytechnic, Aba. e-mail: fman707@yahoo.com
Author ρ: School of General Study, Alvan Ikoku Federal College of Education, Owerri. e-mail: oluwasijibomi@Hotmail.com

In this paper, we propose a set of new techniques for performance measurement in which latency is measured in the context of a workload that is realistic, both in terms of the application used and the rate at which user- initiated events are generated. We present low-overhead methods that require minimal modifications to the system for measuring latency for a broad class of interactive events. We use a collection of simple benchmark examples to characterize our measurement methodology. Finally, we demonstrate the utility of our metrics by applying them in a comparison of Microsoft Windows 9x, Windows 2000, and Windows XP Professional, using realistic interactive input to off-the-shelf applications.

The remainder of this section provides background on the problem of measuring latency, including the motivation for our new methodology based on an analysis of the current practice in performance measurement. Section 2 describes our methodology in detail. In Section 3, we discuss some of the issues in evaluating response time in terms of a user's experience. In Sections 4 and 5, we apply our methodology in a comparison of Windows 9x, Windows 2000, and Windows XP Professional. Sections 6 and 7 discuss the limitations of our work and conclude.

a) The Irrelevance of Throughput

Most macro-benchmarks designed for interactive systems use throughput as the performance metric, measuring the time that the system takes to complete a sequence of user requests. A key feature of throughput as a performance metric is that it can be measured easily, given an accurate timer and a computation that will do a fixed amount of work. Throughput metrics measure system performance for repetitive, synchronous sequences of requests. However, the results of these benchmarks do not correlate directly with user-perceived performance--a critical metric when evaluating interactive system performance. The performance of many modern applications depends on the speed at which the system can respond to an asynchronous stream of independent and diverse events that result from interactive user input or network packet arrival; we call this event handling latency. Throughput metrics are ill-equipped to characterize systems in such ways. More specifically, throughput benchmarks fail to provide enough information for evaluating interactive system

performance and make inappropriate assumptions for measuring interactive systems.

i. *Information Lost*

The results of throughput benchmarks are often reduced to a single number that indicates how long a system took to complete a sequence of events. Although this can provide information about the sum of the latencies for a sequence of events, it does not provide information about the variance in response time, which is an important factor in determining perceived interactive performance.

The insufficient detail provided by throughput benchmarks can also mislead designers trying to identify the bottlenecks of a system. Since throughput benchmarks provide only end-to-end measures of activity, system activity generated by low-latency events cannot be distinguished from that generated by longer-latency events, which have a much greater impact on user-perceived performance. Worse, if such a benchmark includes sufficiently many short-latency events, these short events can contribute significantly to elapsed time, leading designers to optimize parts of the system that have little or no impact on user-perceived performance. In an effort to compare favourably against other systems in throughput benchmarks, designers may even undertake such optimizations knowingly. In this case, bad benchmarking methodology hurts both system designers and end-users.

In addition, user interfaces tend to use features such as blinking cursors and interactive spelling checkers that have (or are intended to have) negligible impact on perceived interactive performance, yet may be responsible for a significant amount of the computation in the over all activity of an application. Throughput measures pro vide no way to distinguish between these features and events that are less frequent but have a significant impact on user-perceived performance.

ii. *Inaccurate User Assumptions*

Throughput benchmarks often drive the system by feeding user input as rapidly as the system can accept it, equivalent to modeling an infinitely fast user. Such an input stream is unrealistic and susceptible to generating misleading results. One of the sources for such errors is batching. Client-server systems such as Windows NT and the X-Window system batch multiple client requests into a single message before sending them to the server. This reduces communication overhead and allows the server to apply optimizations to the request stream, such as removing operations that are overridden by later requests. Although batching improves throughput, it can have a negative effect on the responsiveness of the system.

When a benchmark uses an uninterrupted stream of requests, the system batches requests more aggressively to improve throughput. Measurement results obtained while the system is operating in this mode are meaningless; users will never be able to generate such an input stream and achieve a similar level of batching in actual use. Disabling batching altogether is sometimes possible but does not fully address the problem. An ideal test input should permit a level of batching that is likely to occur in response to real user input.

Overall, throughput measures provide an indirect rather than a direct measure of latency, and as such they can give a distorted view of interactive performance. An ideal benchmarking methodology will drive the system in the same way that real users do and give designers a correct indication as to which parts of the system are responsible for delays or user-perceptible latency. Obtaining such figures requires that we drive the system using an input stream that closely resembles one that an interactive user may generate and more importantly, an ability to measure the latency of individual events.

II. Methodology

Our methodology must provide the ability to measure the latency of individual events that occur while executing realistic interactive workloads. This poses the following set of new challenges:

- Interactive events are short in duration relative to the timer resolution provided by clock APIs in modern operating systems such as Windows and UNIX. Whereas a batch workload might run for millions of timer ticks, many interactive events last less than a single timer interval.
- Under realistic load, there will often be only a fraction of a second between interactive events in which to record results and prepare for the next measurement. Therefore the measurement scheme must have quick turnaround time.
- Perhaps the most challenging problem is collecting the requisite data without access to the source code of the applications or operating system. With source code, it is straightforward to instrument an application to generate timestamps at the beginning and ending points of every interactive event, but this is time consuming at best and not possible given our goal of measuring widely-available commercial software.

Analyzing interactive applications is just as challenging as measuring them. The time during which an application is running can be divided into think time and wait time. Think time is the time during which the user is neither making requests of the system nor waiting for the system to do something. Wait time is the time during which the system is responding to a request for which the user is waiting. Not all wait time is equivalent with respect to the user; wait time intervals shorter than a user's perception are irrelevant. We call

these classes of wait time "unnoticeable." A good example of unnoticeable wait time is the time required to service a keystroke when a user is entering text. Although the system may require a few tens of milliseconds to respond to each keystroke, such small "waits" will be unnoticeable, as even the best typists require approximately 120 ms per keystroke (Ben Shniderman, Designing the user interface, 1992). Distinguishing between wait time and think time is non-trivial, and the quantity and distribution of wait time is what the user perceives as an application's responsiveness. Our measurement methodology must help us recognize the wait time that is likely to irritate users.

In the following sections, we describe the combination of tools and techniques that we use to measure and identify event latency.

a) Experimental Systems

We ran our experiments on a personal computer based on an Intel Premiere III motherboard, with the Intel Neptune chip set and a 650 MHz Pentium processor. Our machine was equipped with a 256KB asynchronous SRAM Level 2 cache, 512 MB of RAM, and a Diamond Stealth 64 DRAM display card. We used a dedicated 10GB Fujitsu disk (model M1606SAU) for each of the operating systems we tested. These disks were connected via a NCR825-based SCSI II host adapter. Both Windows 2000 and Windows XP systems used a NTFS file system, while the Windows 9x system used a FAT32 file system.

b) The Pentium Counters

The Intel Pentium processor has several built-in hardware counters, including one 64-bit cycle counter and two 40-bit configurable event counters as described in Intel Corporation Developers manual (1995). The counters can be configured to count any one of a number of different hardware events (e.g., TLB misses, interrupts, or segment register loads). The Pentium counters make it possible to obtain accurate counts of a broad range of processor events. Although the cycle counter can be accessed in user or system mode, the two event counters can only be read and configured from system mode.

c) Idle Loop Instrumentation

Our first measurement technique uses a simple model of user interaction to measure the duration of interactive events. In an interactive system, the CPU is mostly idle. When an interactive event arrives, the CPU becomes busy and then returns to the idle state when the event- handling is complete. By recording when the processor leaves and returns to an idle state, we can measure the time it takes to handle an interactive event, and the time during which a user might be waiting.

The lack of kernel source code prevents us from instrumenting the kernel to identify the exact times at which the processor leaves or enters the idle loop. Instead, we replace the system's idle loop with our own low-priority process in each of the operating systems. These low- priority processes measure the time to complete a fixed computation: N iterations of a busy-wait loop. The instrumentation code logs the time required by the loop. The pseudo code is as follows:

```
while (space_left_in_the_buffer) {
    for (i = 0; i < N; i++)
        ;
    generate_trace_record;
}
```

We select the value of N such that the inner loop takes 1ms to complete when the processor is idle. In this way we generate one trace record per millisecond of idle time. If the processor is taken away from the idle loop, the loop takes longer than 1ms of elapsed time to complete. Any non-idle time manifests itself as an elongated time interval between two trace records. The larger we make N, the coarser the accuracy of our measurements; the smaller we make N, the finer the resolution of our measurements but the larger the trace buffer required for a given benchmark run.

We wrote and measured a simple microbenchmark to demonstrate and validate this methodology. It uses a program that waits for input from the user and when the input is received, performs some computation, echoes the character to the screen, and then waits for the next input. We measured the time it took to process a key stroke in two ways. First, we used the idle loop method described above to measure the processing time. Figure 1 shows the times at which the samples were collected.

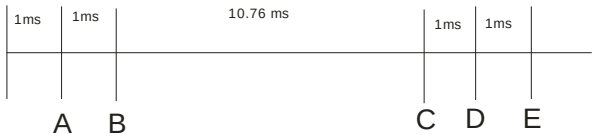


Figure 1 : Validation of Idle loop Methodology. The system spent one ms collecting each of samples A, B, D and E but spent 10.76 ms collecting sample C indicating the system performed 9.76 ms of work during this interval

For the sake of clarity only a few samples are shown. The figure shows that the system spent approximately one ms generating samples A, B, D, and E, indicating that the system was idle during the periods in which these samples were generated, but spent 10.76 ms generating sample C. The difference, (10.76 - 1) or 9.76 ms, represents the time required to handle the event.

Next, we used the traditional approach, recording one timestamp when the program received the character (i.e., after a call to getchar()) and a second timestamp after the character was echoed back to the screen. This measurement reported an event-handling

latency of only 7.42 ms. The 2.34 ms discrepancy between the two measurements highlights a shortcoming of the conventional measurement methodology. Our test program calls the getchar() function to wait for user input. When the user enters a character, the system generates a hardware interrupt, which is first handled by the dynamically linked library KERNEL32.DLL. In the traditional approach, the measurement does not start until control is returned to the test program. Therefore, it fails to capture the system time required to process the interrupt and reschedule the benchmark thread. In comparison, our idle loop methodology provides a more complete measurement of the computation required to process the keyboard event.

Our idle loop methodology uses CPU busy time to rep resent event latency, but there are several issues that pre vent this from being an accurate measure of the user's perceived response time. One problem is that most graphics output devices refresh every 12-17 ms. In this research, we do not consider this effect.

Another problem is that CPU busy time and CPU idle time do not equate directly with wait time and think time. First, synchronous I/O requests contribute to wait time, even though the CPU can be idle during these operations. Second, in the case of background processing, the user may not be waiting even though the CPU is busy. The first problem could be solved with system support for monitoring the I/O queue and distinguishing between synchronous and asynchronous requests. In order to address the second problem, we must consider how events are processed by the systems. When the user generates key strokes and mouse clicks, they are queued in a message queue awaiting processing. Therefore, when there are events queued, we can assume that the user is waiting. By combining CPU status (busy or idle), message queue status (empty or non-empty), and status for outstanding synchronous I/O (busy or idle), we can speculate during which time intervals the user is waiting.

Figure 2 shows a state transition diagram for identifying think time and wait time in our system, using the parameters: CPU state, message queue state, and synchronous I/O status. The diagram omits asynchronous I/O, which we assume is background activity, and assumes that users always wait for the completion of an event. In real ity, we can never precisely distinguish think time from wait time, because we cannot know what the user is doing and whether the user is actually waiting for an event to complete or is thinking while an event is being processed. For simplicity, in the rest of this paper, we assume that the user waits for each event and report results in terms of event handling latency. In the next section, we describe how we obtain information about the status of the message queue.

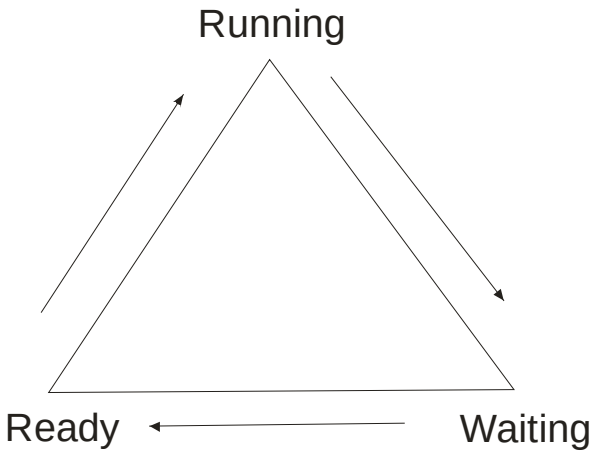


Figure 2 : Showing state transition diagram

d) Monitoring the Message API

Win32 applications use the Peek Message() and Get Message() calls to examine and retrieve events from the message queue. We can monitor use of these API entries by intercepting the USER32.DLL calls. By monitoring use of these API entries, we can detect when an application is prepared to accept a new event and when it actually receives an event. We correlate the trace of Get Message() and Peek Message() calls with our CPU profile to determine when the application begins handling a new request and when it completes a request. This allows us to distinguish between synchronous and asynchronous I/O. It is also useful for recognizing situations where asynchronous computation is used to improve interactive response time.

Figure 2 illustrates our design for a finite state machine that distinguishes think time from wait time in a latency measurement system. In Sections 4, 5, and 6, we will demonstrate how to apply complete information about CPU state and partial information about message queue state to implement part of the FSM. Implementation of the full FSM requires additional system support for monitoring I/O and message queue state transitions. Next, we will present two simple example measurements to give some insight into some of the non-trivial aspects of interpreting the output of our measurements.

e) Idle System Profiles

In this section, we present measurement results for the background activity that occurs during periods of inactivity on Windows 9x and Windows XP. This provides intuition about the measurement techniques as well as baseline information, useful for interpreting latency measurements in realistic situations. Figure 3 shows the idle system profiles for the three test systems. To relate non-idle time to elapsed time, we plot elapsed time on the X-axis and the CPU utilization on the Y-axis. Given that each sample represents 1 ms of idle time, the average CPU utilization during a sample interval can be

calculated easily. For example, if the system spends 10 ms collecting a sample, and the sample includes 1 ms of idle time, the CPU utilization for that time interval is (10 - 1)/10 = 90%.

Both versions of Windows NT show bursts of CPU activity at 10 ms intervals due to hardware clock interrupts. Correlating the samples with a count of hardware interrupts from the Pentium performance counters shows that each burst of computation is accompanied by a hardware interrupt.

Although we have compensated for the overhead introduced by the user-level idle loop, Windows 9x shows a higher level of activity in comparison to both versions of the Windows XP system. We do not know what causes this increased activity in Windows 9x.

By coupling our idle-loop methodology with the Pentium counters, we were able to compute the interrupt handling overhead for various classes of interrupts -- measurements difficult to obtain using conventional methods. For example, the smallest clock interrupt handling overhead under Windows XP was about 800 cycles, or 8 ms.

III. Benchmarks and Metrics

Our benchmark set is organized into three categories. Microbenchmarks are useful for understanding system behaviour for simple interactive operations, such as interrupt handling and user-interface animation. By analyzing microbenchmarks, we develop an understanding of the low-level behaviour of the system. We then extend our measurement to task-oriented benchmarks in order to understand the real impact of latency on the perceived interactive responsiveness of an application. These task- oriented benchmarks are based on applications from typical PC office suites and are designed to represent a realistic interactive computing workload. We further apply application microbenchmarks to evaluate isolated interactive events from the realistic workloads. Our application microbenchmarks include such computations as page-down of a PowerPoint document and editing of an embedded OLE object.

We used Microsoft Visual Test to create most of our microbenchmarks and task-oriented benchmarks. MS Test provides a system for simulating user input events on a Windows system in a repeatable manner. Test scripts can specify the pauses between input events, generating minimal runtime overhead. However, in some cases, the way that Test drives applications alters the behaviour of those applications. This effect is discussed in detail in Section 5.4.

a) Evaluating Response Time

Early in this project, we had planned to develop a new latency metric, a formula that could be used to summarize our measurements and provide a single scalar figure of merit to characterize the interactive performance of a given workload. Events that complete in 0.1 seconds or less are believed to have imperceptible latency and do not contribute to user dissatisfaction, whereas events in the 2-4 second range invariably irritate users Ben Shneiderman (1992). Events that fall between these ranges may be acceptable but can correspond to perceptibly worse performance than events under 0.1 seconds. Our intuition is that a user-responsiveness metric would be a summation of the form:

$$\sum_i f_p(x) \quad \text{where}$$

$$f_p(x) = \begin{cases} 0 & x(i) \le T \\ (x(i) - T)^p & x(i) > T \end{cases}$$

T – user perception threshold
p – some exponent
$x(i)$ – latency of event i

However, we also believe that the threshold, T, is a function of the type of event. For example, users probably expect keystroke event latency to be imperceptible while they may expect that a print command will impose some delay. The issues of event types, user expectation, the precise tolerance of users for delay, and the limitations of human perception are beyond our field of expertise. Presented with these obstacles, we modified our plans, and present latency measurements graphically. We trust that the issues in human-computer interaction can be resolved by specialists. In the meantime, our visualization of latency enables us to compare applications and develop an intuition for responsiveness without risking the inappropriate data reductions that could occur given our limited background in experimental psychology.

IV. Micro-Benchmarks

In this section, we present some basic measurements of simple interactive events. This helps us explore the character of our tools and understand the kinds of things we can and cannot measure. Figure 3

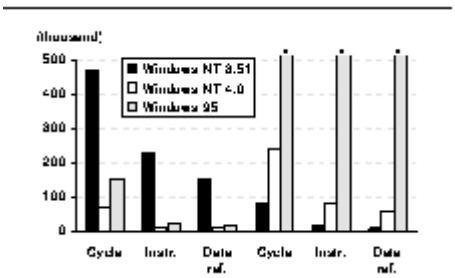


Figure 3 shows the latencies for two simple interactive events, unbound key stroke and mouse click on the screen background, under the three operating systems. We were unable to measure the overhead of

Microsoft Test for these micro-benchmarks, so we were forced to use manual input. To compensate for the potential variability introduced by a human user, we report the mean of 30-40 trials, ignoring cold cache cases. The most significant standard deviations occurred in the key click events for Windows XP and Windows 9x (8%) while all the remaining standard deviations were under 2% of the mean.

On the key stroke test, Windows 9x shows substantially worse performance than Windows XP. This is a reflection of segment register loads (not shown) and other overhead associated with 16-bit and 32–bit windows codes as asserted by Bradley Chen et al, which persist in Windows 9x.

The mouse click results are even more striking. The Windows 9x measurements are off the scale, because the system busy-waits between "mouse down" and "mouse up" events; therefore our measurement indicates the length of time the user took to perform the mouse click. This is much longer than the actual processing times of the Windows XP systems and is not indicative of the actual Windows 9x performance.

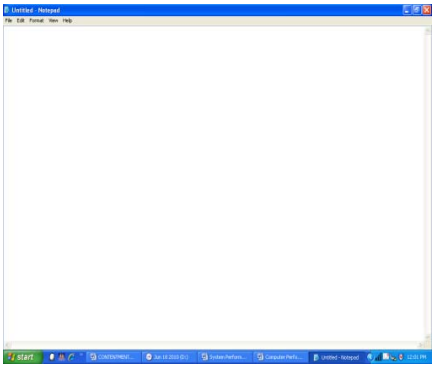
Our methodology provides little guidance in explaining the differences in performance between Windows 2000 and Windows XP Professionals, but it does highlight the fact that instructions and data references occur roughly in pro portion to cycles across the systems for both of the simple interactive events. Therefore, we conclude that in the warm cache case, the performance differences are a function of the code path lengths. It is possible that the difference in code path length stems from the change in GUI between the two notable releases of Windows.

V. Task-Oriented Benchmarks

In this section, we measure three task-oriented bench marks, designed to model realistic tasks that users commonly perform using the target applications. In using these longer running benchmarks we have two specific goals. The first is to measure the system performance for a realistic system state. An often-cited problem of micro-benchmarks is that they tend to measure the sys tem when various caches are already warm. However, measuring the system when all the caches are cold is also unrealistic. Neither extreme is representative of the system state in which the target micro-operations are invoked in common practice. By measuring the latency of micro-operations embedded in a longer realistic inter active task, we measure each micro-operation under more realistic circumstances. The second goal is to identify long-latency operations that users encounter as they perform tasks on the systems. Since these long-latency operations have a greater effect on how users perceive system performance than very short events.

We ran each benchmark five times using Microsoft Test and found that the results were consistent across runs. The standard deviations for the elapsed times and cumulative CPU busy times were 1-2%, and the event latency distributions were virtually identical. The graphical out put shown in the following sections depicts one of the five runs for each benchmark.

a) Microsoft Notepad



Notepad is a simple editor for ASCII text distributed with all versions of Microsoft Windows. Our Notepad benchmark models an editing session on a 56KB text file, which includes text entry of 1300 characters at approximately 100 words per minute, as well as cursor and page movement. With this benchmark, we demonstrate how differences in average response time across the three systems manifest themselves in our visual representation of latency and how they can be used to com pare system performance. We used the same Notepad executable (the Windows XP version) on all three systems and used a Microsoft Test script to drive Notepad. Since virtually all Notepad activity is synchronous, we were able to collect the latency figures for every key stroke that the user made in a straightforward way. By correlating our idle loop measurement with our monitoring of the Peek Message() and Get Message() API calls, we were able to clearly identify the Test overhead and remove it from the data presented.

The cumulative latency graph shows that for all three systems, over 80% of the latency of Notepad is due to low-latency (less than 10 ms) events. These short-latency events are the keystrokes that generate printable ASCII characters. The remaining 20% of the total latency are due to the longer latency (at least 28 ms) keystrokes that cause "page down" or newline operations. These keystrokes cause Notepad to refresh all or part of the screen. Events of the same type contribute equally to the total latency.

The latencies measured are relatively small for Notepad and reflect both the simplicity of the application and the relatively fast PC that we used for our experiments. Although these differences in latency are likely to go unnoticed by users of our test system, they

might have a significant effect on user-perceived performance on a slower machine.

b) Microsoft Notepad

PowerPoint, from the Microsoft Office suite, is a popular application for creating presentation graphics. In our PowerPoint task scenario, the user starts PowerPoint immediately after powering up the machine and booting the operating system, so that all caches are cold. The user then loads a 46-page, 530KB presentation, and finds and modifies three OLE embedded Excel graph objects. Each of the OLE objects was of similar size and complexity. As with Notepad, we used a Microsoft Test script to drive the application and deliver key strokes at a realistic rate, with each keystroke separated by at least 150 ms. An important property of the PowerPoint benchmark is that it has a number of events with easily perceptible latencies. Since we were mainly interested in longer events, we pre-processed our data to exclude events with latency of less than 50 ms. Figure below shows the results for the two versions of Windows. We were unable to run this experiment for Windows XP due to limitations of Microsoft Test when manipulating OLE embedded object on that system.

The shortest event (with latency of less than one second), are due to "page down" operations and MS-Excel operations. Both systems exhibited a similar latency distribution for these events. Six events had latencies greater than one second on both systems, in nearly the same relative order. Table 1 lists these long latency events.

All of the long-latency events required disk accesses, which are responsible for the majority of the latency for these events. The effects of the file system cache are most clearly observed in the latency for starting the second OLE edit, as more of the pages for the embedded Excel object editor become resident in the buffer cache.

The cumulative latency graph shows that both versions of Windows 2000 and Windows XP demonstrate similar performance for the short-latency keystrokes, and the majority of the performance difference is a result of the ability of NTFS file system to handle the long-latency events much more efficiently.

	Latency (in seconds)	
	NT 3.51	NT 4.0
Save document	6.062	9.560
Start Powerpoint	7.166	5.773
Start OLE edit session (first time)	7.050	5.644
Open document	5.660	4.151
Start OLE edit session (second object)	2.697	2.009
Start OLE edit session (third object)	2.697	1.305

Table 1.

The standard deviations are all below 3%

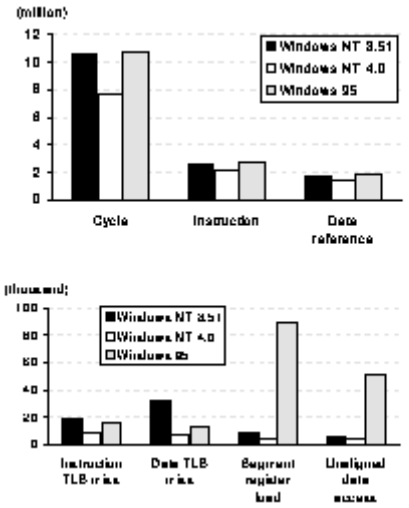


Figure 4 : Latency for simple interactive events

c) Microsoft Word

Our task-oriented workload for Microsoft Word consists of text entry of a paragraph of approximately 1000 characters. It includes cursor movement with arrow keys and backspace characters to correct typing errors. The timing between keystrokes was varied to simulate realistic pauses when composing a document, and line justification and interactive spell checking were enabled. We do not report results for Windows 9x, because the system does not become idle immediately after Word finishes handling an event, making all event latencies appear to be several seconds long.

Figure 11 shows results for Microsoft Test driven simulations on the two versions of NTFS based Windows. Compared to Notepad, MS-Word requires substantially more processing time per keystroke, due to additional functionality such as text formatting, variable-width fonts and inter active spell checking. For the majority of interactive events, Windows XP exhibits shorter response time and lower variance than Windows 2000.

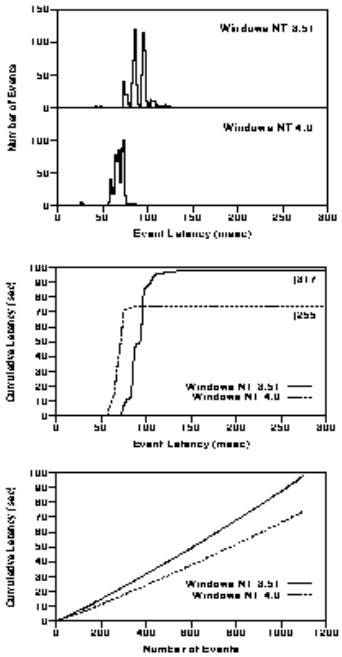


Figure 5 : Notepad Event Latency Summary

The MS-Word benchmark demonstrates both the strengths and limitations of evaluating interactive performance using latency. Compared to throughput measurements, our latency analysis provides much more detailed information, such as variations in latency and the distribution of events with different latencies. How ever, the structural features of Word push us to the limit of the behavior we are able to analyze. Our analysis indicates that Word uses a single system thread, but responds to input events and handles background computations asynchronously using an internal system of coroutines or user level threads.

Distinguishing background activity from foreground activity in MS-Word is challenging. We examined the results of hand-generated Word input under Windows 2000 OS, compared it to the Test-generated results, and found significant differences. For our hand-generated tests, we ran seven trials, with the same typist and input, and found that the event histograms appeared very similar and that the variation in cumulative latency and elapsed time was less than 4% across the runs. While the Test results showed that most events had latency between 80 and 100 ms, we measured a 32 ms typical latency for the hand-generated input. This difference in event latency was accompanied by a compensating difference in back ground activity. The hand-generated input showed a higher level of background activity than the Test-generated results. We also observed that carriage returns under the hand-generated input took longer than 200 ms to handle while the longest latency events we saw in the Test-generated runs were 140 ms. Our Message API log reveals that Test generates a WM_QUEUESYNC messages after every keystroke. We hypothesize that these messages were responsible for the different behavior under Test and under manual typing. However, with our current tools, the complexity of Word makes it difficult to thoroughly analyze even the simple experiment we present here.

VI. Summary

The tools and techniques we have discussed here are a first step towards understanding and quantifying interactive latency, but there remains much work to be done. In the absence of system and application source, better performance monitoring tools would be useful. Our measurements could be improved through API calls that return information about system state such as message queue lengths, I/O queue length, and the types of requests on the I/O queue. Currently, some of this information can be obtained, but it is painful (e.g., monitoring the Get Message() and Peek Message() calls).

Even in the presence of rich APIs, the task of distinguishing between wait time and think time is not always possible. There is no automatic way to detect exactly what a user is doing. Without user input, we can never tell whether a user is genuinely waiting while the system paints a complicated graphic on the screen or is busy thinking. For simulations using designed scripts, we can make assumptions about when users think and then analyze performance based on those assumptions, but the most useful analysis will come from evaluating actual user interaction.

One factor that contributes to user dissatisfaction is the frequency of long-latency events. We processed the Microsoft Word profile of Figure 5 to analyze the distribution of inter-arrival times of events above a given threshold. Since most events in the Word benchmark were very short, we chose thresholds around 100 ms. Table 2 shows the summaries for these thresholds. Note that the standard deviations are of the same order of magnitude as the averages themselves, indicating that there is no strong periodicity between long-latency events.

Threshold in ms	No of Events above Threshold	Inter-arrival times	
		Average (in sec)	Std Deviation (in sec)
100	101	3.1	3.1
110	26	12.4	10.6
120	8	41.1	48.8

Table 2

We then examined the truly long-latency events from the PowerPoint benchmark. Figure 12 shows the event latency profile for all events over 50 ms. Both systems show similar periodicity with the better performing 4.0 system demonstrating smaller inter-arrival times to match its shorter overall latency.

In the case of Word, the inter-arrival times are clustered because most events have similar latency. In the case of PowerPoint, the inter-arrival times of long-latency events are simply the inter-arrival times of a few particular classes of events. The distribution of these events is entirely dependent upon when we issued such requests in our test script and is not necessarily indicative of the distribution that might be obtained from a real user. In this test, none of the simple keystroke events were responsible for generating long-latency events, rather all the events with latencies over 50 ms result from major operations for which user expectation for response time is generally longer. Until our tools become sophisticated enough to examine long traces of complex events generated by a real user, further analysis of these inter-arrival times is not particularly productive.

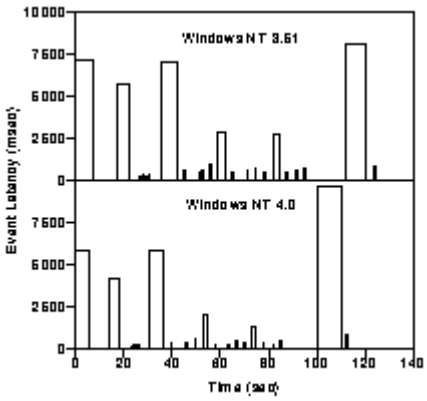


Figure 6 : Counter measurements for PowerPoint page down operation

Over time, our tools will become better able to deal with the sophisticated applications that we seek to analyze, but we need the human factors community to assist us in understanding the limits of human perception and the models of user tolerance. Some of the questions that must be answered are:

- What are the limits of human perception?
- How do the limits vary by task (e.g., typing versus mouse-tracking)?
- How do the user expectation and tolerance for interactive response time vary by task?
- How does user dissatisfaction grow with increasing of latency?
- How does user dissatisfaction grow with the variance of latency?
- What aspects of performance contribute the most to user satisfaction?

VII. Conclusions

Latency, not throughput, is the key performance metric for interactive software systems. In this paper, we have introduced some tools and techniques for quantifying latency for a general class of realistic interactive application. To demonstrate our methodology, we applied it to compare the responsiveness of realistic applications running on three popular PC operating systems. Whereas current measurements of latency are generally limited to micro-benchmarks, our approach allows us to measure latency for isolated events in the context of realistic interactive tasks. Our latency measurements give a more accurate and complete picture of interactive performance than throughput measurements.

We have combined a few simple ideas to get precise information about latency in interactive programs. We have shown that using these ideas we can get accurate and meaningful information for simple applications and also, to a degree, for complex applications. The requirements of these techniques are not out of reach; in particular, a hardware cycle counter, a means for changing the system idle loop, and a mechanism for logging calls to system API routines are needed. Additional support for detecting the enqueuing of messages and the state of the I/O queue would provide a more complete framework for latency measurement. We have shown the limitations of our system for applications such as Microsoft Word that use batching and asynchronous computation.

Measuring latency for an arbitrary task and an arbitrary application remains a difficult problem. Our experience with Microsoft Word demonstrates that there are many difficult technical issues to be resolved before latency will become a practical metric for system design. Our graphical representation provides a great deal of information about program behavior to specialists, but is probably not appropriate for more widespread use. The two key components necessary to provide consumers a single figure of merit are further work in human factors and some method for distinguishing user think time from user wait time.

VIII. Acknowledgments

The authors have benefited from the works of Brian N. Bershad, et al, John K. Ousterhout, and Jeffrey C. Mogul of Business Applications Performance Corporation. We thank them for their insights and suggestions. The authors are greatly grateful and indebted to our research partners and colleagues in the School of Science, and Computer Engineering, Abia State Polytechnic, Aba, who have being research fellows and have supported our research for some years now. We would also like to thank the Management of both Abia State Polytechnic, Aba, and Alvan Ikoku Federal College of Education, Owerri for providing the platform and the right atmosphere for this work.

References Références Referencias

1. Business Applications Performance Corporation, "Sysmark for Windows NT," Press Release by IDEAS International, Santa Clara, CA, March 1995.
2. Brian N. Bershad, Richard P. Draves, and Alessan dro Forin, "Using Microbenchmarks to Evaluate System Performance." Proceedings of the Third Workshop on Workstation Operating Systems, IEEE, Key Biscayne, Florida, April 1992, pages 148-153.
3. Ben Smith, "Ultrafast Ultrasparcs," Byte Magazine, January 1996, page 139. Additional information on the Bytemarks suite is available on the Internet: http://www.byte.com/bmark/bdoc.htm.
4. J. Bradley Chen, Yasuhiro Endo, Kee Chan, David Mazieres, Antonio Dias, Margo Seltzer, and Michael D. Smith, "The Measured Performance of Personal Computer Operating Systems," ACM Transactions on Computer Systems 14, 1, February 1996, pages 3-40.
5. Intel Corporation, Pentium Processor Family Developer's Manual. Volume 3: Architecture and Programming Manual, Intel Corporation, 1995.
6. C. J. Lindblad and D. L. Tennenhouse, "The VuSystem: A Programming System for Compute-Intensive Multimedia," To appear in IEEE Journal of Selected Areas in Communication," 1996.
7. Larry McVoy, "Lmbench: Portable tools for performance analysis," Proceedings of the 1996 USENIX Technical Conference, January 1996, pages 179- 294.
8. Jeffrey C. Mogul, "SPECmarks are leading us astray," Proceedings of the Third Workshop on Workstation Operating Systems, IEEE, Key Biscayne, Florida, April 1992, pages 160-161.
9. James O'Toole, Scott Nettles, and David Gifford, "Concurrent Compacting Garbage Collection," The Proceedings of the Fourteenth ACM Symposium on Operating System Principles, December 1993, pages 161-174.
10. John K. Ousterhout, "Why Operating Systems Aren't Getting Faster As Fast As Hardware." Proceedings of the Summer 1991 USENIX Conference, June 1991, pages 247-256.
11. Mark Shand, "Measuring Unix Kernel Performance with Reprogammable Hardware," Digital Paris Research Lab, Research Report #19, August 1992.
12. Ben Shneiderman, Designing the User Interface, Addison-Wesley, 1992.
13. Jeff Reilly, "SPEC Discusses the History and Reasoning behind SPEC 95," SPEC Newsletter, 7(3):1- 3, September 1995.
14. M. L. VanNamee and B. Catchings, "Reaching New Heights in Benchmark Testing," PC Magazine, 13 December 1994, pages 327-332. Further information on Ziff-David benchmarks is available on the Internet: http://www.zdnet.com/zdbop/.

Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 14 Issue 5 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 Handoff Latency in Advanced Wireless Networks

By B. Jaiganesh
Saveetha University, India

Abstract- The association of different wireless communication technologies on the way to advanced wireless networks had better face with the developing systems resource utilization and user authentication. Mobility management is vital to omnipresent computing which can be established by location management and distinctive of the mobility management modules. In this work the new protocol is proposed which includes the integration of FHMIPv6 and MIH. The proposed protocol performance is analysed using NS2 simulation. It shows the reduction of handoff latency for video streaming. The cost is also being reduced by the handoff latency while transmitting the signal from one mobile user to another. Further the proposed protocol is compared with the previous protocols.

Keywords: handoff latency, 4g wireless web, flexibility management, handoff progression and situation management, NS2, FHMIPV6– MIH proposed integrated solution.

GJCST-G Classification: C.2.1

Strictly as per the compliance and regulations of:

Analysis of Handoff Latency in Advanced Wireless Networks

B. Jaiganesh

Abstract- The association of different wireless communication technologies on the way to advanced wireless networks had better face with the developing systems resource utilization and user authentication. Mobility management is vital to omnipresent computing which can be established by location management and distinctive of the mobility management modules. In this work the new protocol is proposed which includes the integration of FHMIPv6 and MIH. The proposed protocol performance is analysed using NS2 simulation. It shows the reduction of handoff latency for video streaming. The cost is also being reduced by the handoff latency while transmitting the signal from one mobile user to another. Further the proposed protocol is compared with the previous protocols.

Keywords: handoff latency, 4g wireless web, flexibility management, handoff progression and situation management, NS2, FHMIPV6– MIH proposed integrated solution.

I. Introduction

The federation of different wireless communication technologies on the way to 4G wireless networks had better face some anticipated challenges in advance representative practice implementation. One of the major challenges is the mobile station mobility managing by dissimilar wireless technologies in mandate to acquire the mobile station linked to the unsurpassed available wireless network. To amalgamate these perpendicular wireless networks in one network as a triggered network that can be acquiesce an improved service at lower cost to the manipulator, as well as progress the overall networks resource consumption. However, accomplishing these two goals needs an elegant mobility management system that can be achieved the trade-off flanked by efficient resource utilization and mobile station grasped QoS. Mobility management excludes two parts, handoff and location management. As soon as a mobile station moving across the boundary of dualistic neighbour cells, the MSC prepares a innovative twofold channels in the fresh cell to conserve the call commencing dropping, this operation is called a Handoff Management (HM). The location management (LM) is pursuing the active mobile station (powered on MS) while roaming without a call. Despite the fact the location of a MS essential be known accurately during a call, LM habitually means in what way to trackan active mobile station between two sequential phone calls. The peak important issues in mobility management are seamless roaming (integration among different 4G wireless networks, QoS assurance, operational costs buoyed features and a good utilization of the wireless links (utilizing the wireless acquaintances represented by inhabiting the rheostat channels in the bleeping and location apprise operations). Additionally, perpendicular handoff flanked by radio admittance networks consuming poles apart technologies entail additional adjournment for relinking the mobile terminal to the innovative wireless access network, which may foundation packet losses and degrade the QoS for concurrent traffic. The habitation of bandwidth, entirely computational processes in substructure of the network, power ingestion in MS, plus power consumption in the network are form the cost and all of this is a commercial cost. Therefore the cost bargain is a appropriate important issue in LM. The intentions of this paper are to single-mindedness on handoff management (HM), which is an vital component of mobility management, in aiding seamless mobility across heterogeneous network infrastructures. Correspondingly focusing on the altered protocols in handoff management and equate those protocols for audio, video & FTP (file transfer protocol) transmission.

II. MIPV6Protocol

When the surroundings change, the Mobile IPv6 protocol permits mobile nodes to access IP address sub network to continue communications with the communication on the side. Mobile IPv6 (C. Perkins *et al.,* 2004) architecture is contained of three key elements: a Mobile node (MN), Home Agent (HA), Correspondent Node (CN). The main processes of Mobile IPv6 are:

1. The regular route of communication is followed by the Mobile Node when it is linked to its home agent link.
2. The neighbour discovery (ND) device to discover whether itself has roaming on a foreign agent link via the Mobile Node.
3. It will obtain Care of Address (CoA) on the foreign agent connection through the address auto configuration procedure, when the Mobile Node has found itself to travel to the field on the link, on the base of the access router declaration facts.
4. It can retain the earlier CoA, and login on the home agent CoA recognized as the primary Care of

Author: Saveetha University, Chennai.
e-mail: jaiganesh.price@gmail.com,

Address, and the Mobile Node to its CoA through the binding update information logs on to the home agent.

5. This Mobile Node informs its communicating on the client its CoA to the basis of make sure the protection.
6. When the mobile node side does not know its CoA, its HA link will interrupt these packets and then use method to forward those packets to the Mobile Node. It will send the information packet from its home network clearance to its home address.
7. It uses IPv6 routing header to direct packets to the Mobile Node, when the announcement to the client recognizes the Mobile Node CoA.
8. When the Mobile Node obtains the packet and recognizes it to be forwarded by the Home Agent link, it informs the CoA to the source node of this packet so that the source node can afterwards be under the CoA packets sent directly to the Mobile Node, and the home agent(HA) link no longer shall forward.
9. It forwards the packet via the Mobile Node through the tunnel, as per the binding update information which is identified by it, when the Mobile Node is on the connection, where the earlier default router obtains a packet which is sent to the Mobile Node. In this point the role of the default router is related to the Mobile Node's Home Agent, when the Mobile Node to communicate with other nodes in the other direction. The message packet uses a special method to be routed directly to the destination. If there is a robust security requirement the Mobile Node uses the tunnel to send the information to the Home Agent, and then sent by the Home Agent to the primary address of the tunnel for the Mobile Node's Care of Address.

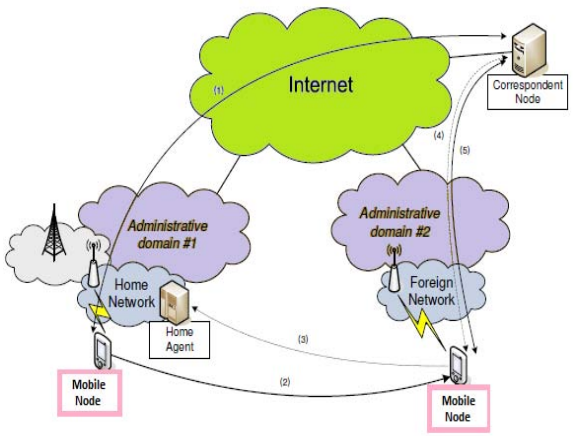


Figure 1 : Mobile IPv6

III. HMIPv6 Protcol

The modification of the basic MIPv6 protocol in the binding registration procedure in the Hierarchical Management is shown through the introduction of location management mechanism, and decrease the registration frequency of the Mobile Node to the remote CN and HA for decreasing HO latency. Now a days the level switch system Hierarchical MIPv6 (Soliman.H *et al.*,2005) becomes a standard hierarchical management class switch program. Compared to Fast Handover for MIPv6 (FMIPv6) the entire performance is better in Hierarchical MIPv6. Different kinds of methods are proposed based on different features. For the case in point Care of Address (CoA) group established on Hierarchical Mobile IPv6 HO system, the foremost idea of this system is to introduce address pool in the access router and MAP (Mobile Anchor Point), the address used in this wireless network is kept in the address group, removing the essential for Care of Address (CoA) of the Duplicate Address Detection (DAD) process.

This arrangement expands MAP protocol of the Hierarchical MIPv6 in Mobile Anchor Point discovery protocol in Hierarchical management, completes the function that the Mobile Node takes the Mobile Anchor Point agency logically and chooses the Mobile Anchor Point discovery protocol on the router to create apparent, which is easy to support. The Mobile Anchor Point discovery protocol's benefits over Mobile Anchor Point discovery protocol in Hierarchical MIPv6 is that the mobile node can wisely choose the Mobile Anchor Point support and create Mobile Anchor Point discovery protocol apparent to the router, so that the protocol is easy to uphold. The disadvantage that the mutual swapping information between Mobile Anchor Point agents is desirable to preserve Mobile Anchor Point topology table.

The mobile node needs to use the new algorithm for finding the nearby Mobile Anchor Point agent. It increases the interface load of the region signalling and the design complexity of Mobile Anchor Point agent and mobile node. The Hierarchical MIPv6 is using the sorting management programs of Mobile Anchor Point discovery protocol; it's similar to MIPv6, but it is complex than MIPv6.

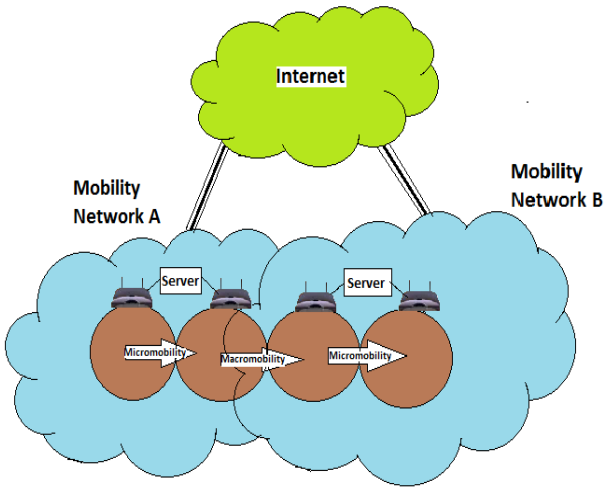


Figure 2 : Hierarchical MIPv6

IV. PMIPV6 Protocol

Proxy Mobile IPv6 (or PMIPv6, or PMIP) is a network-based mobility management protocol homogenous by IETF is a protocol for edifice a common and entree technology sovereign of mobile core networks, accommodating various entree technologies such as WiMAX, 3GPP, 3GPP2 and WLAN based access architectures. Proxy Mobile IPv6 is the merely network-based mobility management protocol standardized by IETF.

Advantages

- Handover performance optimization: PMIPv6 can condense the latency in IP handovers by preventive the mobility management within the PMIPv6 domain. Therefore, it can largely avoid remote service which not only cause long service delays but consume more network resource.
- Reduction in handover-related signaling overhead. The handover-related signaling overhead can be aggravated inPMIPv6 since it avoids tunnelling overhead over the air and as well as the remote Binding Updates either to the Home Agent (HA) or to the Correspondent Node (CN).
- Location privacy. Keeping the mobile node's Home Address (MN-HoA) unchanged over the PMIPv6 domain dramatically condenses the chance that the attacker can construe the precise location of the mobile node.

Applications

- Selective IP Traffic Offload Support with Proxy Mobile IPv6.
- Network-based Mobility Management in a local domain (Single Access Technology Domain).
- Inter-technology handoff across access technology domains (Ex: LTE to WLAN, eHRPD to LTE, WiMAX to LTE).
- Access Aggregation replacing L2TP, Static GRE, CAPWAP based architectures, for 3G/4G integration and mobility.

Network based mobility management enables the same functionality as MIP, wihout any changes in the host TCP/IP protocol stack by PMIPV6 the host can change its point of attachment to the Internet without changing its IP address.

PMIPv6 is transparent to mobile nodes, PMIPv6 is used in localized networks with limited topology where handover signalling delays are minimal.

V. FMIPV6 Protocol

The advantage of some programs is that FMIPv6 efficiently decreases HO latency and Packet loss of the performance is improved in Fast handover scheme (Rajeev Koodli 2004), such as presenting link layer mobility calculation or link layer trigger methods, new CoA configuration, and duplicate address detection (DAD) procedure.

The old router will obtain a request broker news RtSol from the NAR, the necessity to go into a new subnetwork, when the Mobile Node as the second level activate being conscious. The NAR well along proceeds cut start news Handoff Initiate (HI) obtains from the old router. Then it sends a verification message HACK after receiving the message from NAR. The old access router sends a Router Advertisement (RtAdv) message to MN as an agent on a router solicitation message reply, and Mobile Node gets the CoA.

Router Advertisement message directed by the old router is received by the Mobile Node and Mobile Node gets a F-BACK (fast binding acknowledgment message), and to the network in which the old router is positioned and to the (NAR) NAR network through the tunnel. It has worked with a new subnetwork conventional after the second layer link. When the Mobile Node gets to a new sub network, a fast neighbour advertisement message F-NA is issued by the Mobile Node, and then (NAR) new access router can forward message to Mobile Node. It can be found that in feature of handoff delay after thorough investigation of the handoff process, the mobile monitoring. The Fast Handover for MIPv6 (FMIPv6) protocol eliminates the basic mobile IPv6 HO procedure, the duplicate address detection (DAD) delay and new Care of Address (CoA) configuration.

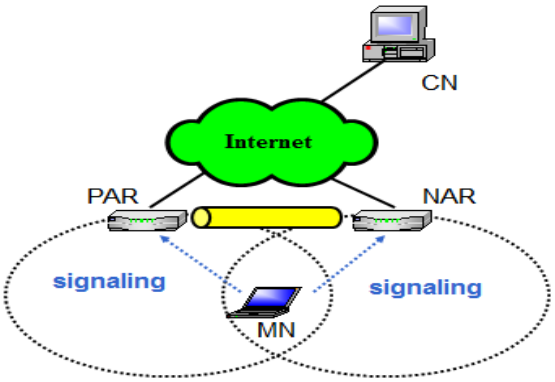


Figure 3 : FMIPv6

VI. FHIMPV6 Protocol

The important handoff management parameters are to enhance and achieve the HO delay and packet loss. Present days , its more broad application of such programs, MIPv6 application layer management structure use the fast handoff system, which effectually links the fast handoff scheme and hierarchical management program that Fast Handover Support in Hierarchical Mobile IPv6 (H. Y. Jung *et al.,* 2005) (FHMIPv6), and shows good handoff presentation.

The FMIPv6 and HMIPv6 is applied the both in the main principle of FHMIPv6, the Mobile IPv6 (MIPv6) protocol at the same time is not a simple arrangement of the two, it will cause triangular routing problem. The previous access router (PAR) through MAP agent that the data packet sent to Mobile Node will be carried. Then convey the packet to NAR to the previous access router (PAR), in the hierarchical network topologies, forming a triangle routing, the data packet will go through the Mobile Anchor Point agent once more.

The optimization of data flow is realized in Fast Handover Support when Hierarchical Mobile IPv6 (H. Y. Jung *et al.,* 2005) selects Mobile Anchor Point agent as an alternative to Previous access router. In other than pass the previous access router, which the data packet sent to the Mobile Node, is sent to new access router (NAR) openly through Mobile Anchor Point agent, for escape the triangle routing. The request message to Mobile Anchor Point is to get the new forward address from the Mobile Node sending a router agent. Mobile Anchor Point will coming back a router agent declaration to Mobile Node as soon as it obtains the message then Mobile Node will form a new transfer address and direct bring up-to-date information about the fast binding to Mobile Anchor Point.

The Mobile Anchor Point starts the handoff procedure between the access routers through a primary message to the new access router after receiving it. The handoff initial message is obtained by the new access router, notices proficiency of the new forward address, and Mobile Anchor Point is getting the acknowledged information. The NAR and the Mobile Anchor Point are set up to make the two-way tunnel between them. Mobile Anchor Point sends an acknowledged message of fast binding to Mobile Node, after getting the information. It sends efficient fast binding information to the NAR, as soon as Mobile Node knows the link information. The NAR then transports data to Mobile Node from the above handoff procedure.

The features of reducing the HO delay, and Packet loss, also evades the triangle routing problem, that the fact that Fast Handover Support in Hierarchical Mobile IPv6 links the advantages of Fast Handover for MIPv6 and Hierarchical MIPv6 works very well. But growths the complexity of designing a Mobile Anchor Point agent and the problem of Mobile Anchor Point agent.

VII. FHMIPv6– MIH Proposed Integrated Solution

The network based mobility management solution in the simulation of mobility across coinciding wireless access networks in micro mobility domain in the simulation setup was implemented. The integrated solution proposed setup is the same as the FHMIPV6 and integrates IEEE802.21 functionality in the MN and the ARs.

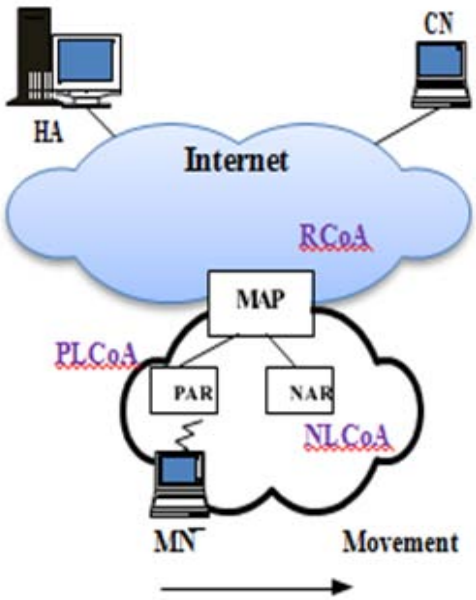


Figure 4 : FHMIPV6 – MIH Proposed integrated solution

VIII. Simulation Setup

This simulation shows that the PAR and NAR are in isolated sub networks. The two ARs have both Data Link Layer and Network Layer abilities that grips HOs. They are organized in a hierarchical tree structure of point-to-point wired links, and the router is interrelated to the MAP by a series of agents.

The MN to the CN using my UDP to the stream of video traffic is simulated and transmitted. The video packet size is established at 1028 B while the break among successive packets is also stable at 1ms.

Thus the Figure 5 shows simulation setup FHMIPV6-MIH Proposed integrated solution of using NS-2 Simulation Setup.

Both CN and HA are connected to an intermediate node (AR1) with 2ms link delay and 100 Mbps links. The link between AR1 and the MAP is a 100 Mbps link with 50 msec link delay. The MAP is further connected to the intermediate nodes AR2 and AR3 with 2 msec link delay over 10 Mbps links. AR1 and AR2 are connected to PAR and NAR with 2 msec link delay over 1 Mbps links.

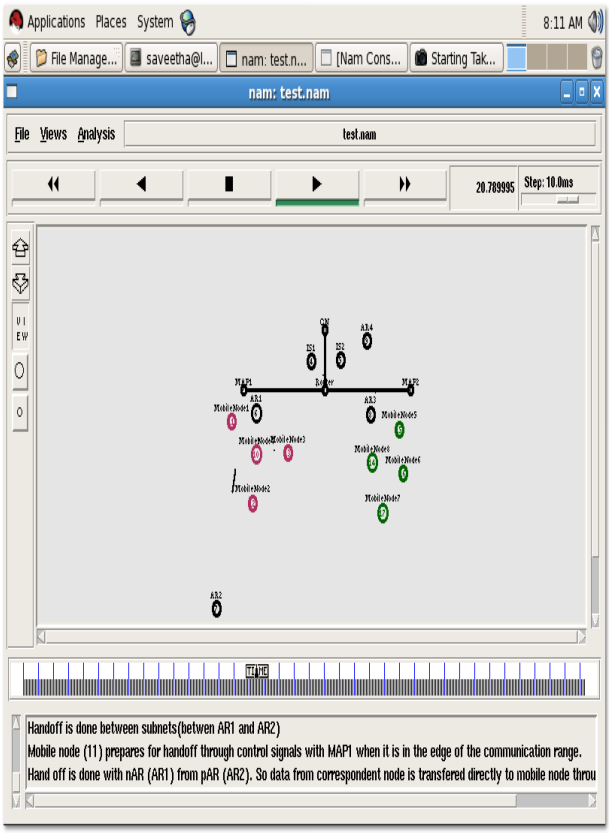


Figure 5: FHMIPV6-MIH Proposed integrated solution of using NS-2 Simulation Setup

IX. Results and Discussion

Simulation results are obtained as follows:

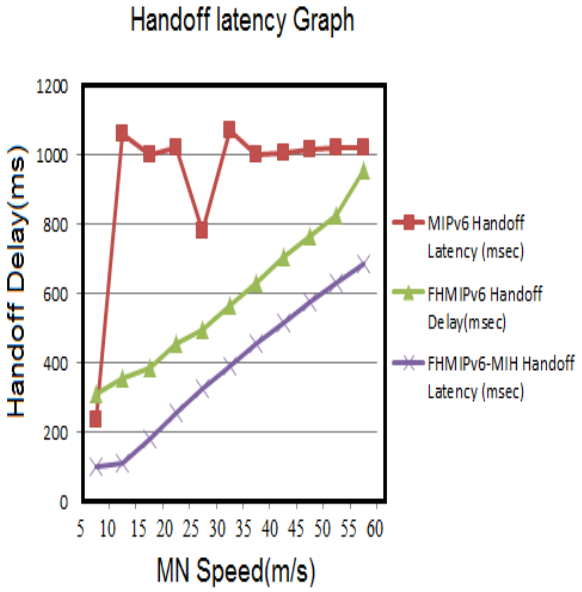


Figure 6 : HO latency Graph

Figure 6 shows the Handoff delay for MIPv6, FHMIPv6 and FHMIPv6-MIH scenarios gained during the Simulation. It shows handoff delay for MIPv6 is red line and the green line specifies delay produced with FHMIPv6 and the blue line shows the proposed method of FHMIPv6-MIH is blue line. Three seconds into the simulation, when the Mobile Node starts moving, MIPv6's handoff delay arises to increase peaking at 8 seconds with 1000 msec. The delay remains at 1000 msec up to the end of the simulation except at twenty one seconds when delay decreases to 790 msec. In contrast, FHMIPv6's delay is at average 586msec. But this delay made the interruption between the Mobile nodes. Then the proposed method, average handoff latency is at 384 msec when horizontal handover takes place. Figure 6 proves that FHMIPv6-MIH practices less latency than MIPv6 and FHMIPv6. Less latency shows that communication between the Mobile Node and the Correspondence Node will have an improved quality in communication.

Comparative Analysis of different protocols of Handoff Latency tabulation 1:

MN Speed(m/s)	MIPv6 Handoff Latency (msec)	FHMIPv6 Handoff Delay(msec)	FHMIPv6 MIH Handoff Latency (msec)
5	236.25	310.24	102.45
10	1062.38	355.57	109.02
15	1000.87	388.14	183.05
20	1020.95	455.36	254.60
25	780.85	495.79	323.64
30	1070.23	564.84	390.20
35	1000.57	632.12	454.25
40	1005.14	708.08	515.82
45	1015.12	763.78	574.88
50	1020.32	824.83	631.46
55	1022.65	956.11	685.53

Comparative Analysis of different protocols of Handoff Latency in FTP, Audio and Video tabulation 2:

PROTOCOL	HANDOFF LATENCY (msec)		
	FTP	AUDIO	VIDEO
MIPV6	5487	(50-250)	(100-300)
HMIPV6	739	400	(300-500)
FMIPV6	532	-	200
FHMIPV6	301	-	(200-400)
PMIPV6	-	-	406
FHMIPV6& MIH	-	-	120

X. Conclusion

In this paper mobility management has been enhanced in 4G especially in Handoff Management. On compression with the results using various Network layer protocols such MIPV6, HMIPV6, FMIPV6, FHMIPV6, PMIPV6, & FHMIPV6-MIH. The proposed FHMIPV6-MIH protocol yields better results. Due to the tendency of fast mobile user having the coverage area is high. The velocity is increased and also the cost is reduced due to the handoff latency while transmission of signal from one mobile user to another. By the comparative analysis of different protocols, the handoff latency of video is drastically reduced in FHMIPV6-MIH to 120msec, which can be used for future applications. These simulation results show that as the velocity increases, the number of handoff will also increases. This scenario happened because of the tendency of fast mobile user to leave the coverage area is high compared to slow mobile user. Therefore, the number of handoff is increasing with reverence to the velocity of the mobile user. The cost is also be decreased due to the handoff latency while transmitting the signal from one mobile user to another mobile user.

References Références Referencias

1. Yang Jie and DheyaaJasimKadhim "Performance evaluation of the mobility management towards 4g wireless networks " volume 4, issue 5, september – october, 2013.
2. I.F. Akyildiz, Jiang Xie and S. Mohanty, "A survey of mobility management in next- generation all-IP-based wireless systems", IEEE Wireless Communications Journal, Vol. 11 No. 4, 2004.
3. Robert Hsieh, Aruna Seneviratne1, HeshamSoliman, Karim El-Malki "Performance analysis on Hierarchical Mobile IPv6 with Fast-handoff over End-to-End TCP" .
4. ZimaniChitedze, Z., & Tucker, W. D. (2012). FHMIPv6-based handover for wireless mesh networks. In S. Scriba (ed.), Southern Africa Telecommunication Networks and Applications Conference (SATNAC 2012), pp. 135–139, George, South Africa.
5. Shaojian Fu, Mohammed Atiquzzaman "Handover latency comparison of SIGMA, FMIPv6,HMIPv6, and FHMIPv6".
6. Liu Jing, Jungwook Song, Heemin Kim, Boyoung Rhee, SunyoungHan"Multi-SessionMethod to Reduce Handover Latency and Data Losing in FHMIPv6"
7. Koodli R. (editor) (2004), "Fast handovers for Mobile IPv6," IETF DRAFT, draft-ietf-mipshop-fast-mipv6-03.txt.
8. Hui, S.Y. and K.H. Yeung, Challenges in the migration to 4G mobile systems. Communications Magazine, IEEE, 2003. 41(12): p. 54-59.
9. Jung H.Y., Koh, S.J. and Lee, J.Y. (2005), "Fast Handover for Hierarchical MIPv6 (FHMIPv6)", IETF Internet Draft draft-jung-mobopts-fhmipv6-00.txt.
10. Akyildiz, I.F., J. Xie, and S. Mohanty, A survey of mobility management in next-generation all-IP-based wireless systems. Wireless Communications, IEEE, 2004. 11(4): p. 16-28.
11. Perkins, C.E., Mobile IP. Communications Magazine, IEEE, 2002. 40(5): p. 66-82.
12. Lee H., Han Y. and Min S. (2010) "Network Mobility Support Scheme on PMIPv6 Networks", International Journal of Computer Networks and Communications (IJCNC), Vol.2, No.5, pp. 19-23.
13. H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier, "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)," IETF RFC 4140, Aug. 2005.

Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 14 Issue 5 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 Text Mining-based Anomaly Detection Model in Network Security

By Mohsen Kakavand, Norwati Mustapha, Aida Mustapha & MohdTaufik Abdullah
Putra University, Malaysia

Abstract- Anomaly detection systems are extensively used security tools to detect cyber-threats and attack activities in computer systems and networks. In this paper, we present Text Mining-Based Anomaly Detection (TMAD) model. We discuss n-gram text categorization and focus our attention on a main contribution of method TF-IDF (Term frequency, inverse document frequency), which enhance the performance commonly term weighting schemes are used, where the weights reflect the importance of a word in a specific document of the considered collection. Mahalanobis Distances Map (MDM) and Support Vector Machine (SVM) are used to discover hidden correlations between the features and among the packet payloads. Experiments have been accomplished to estimate the performance of TMAD against ISCX dataset 2012 intrusion detection evaluation dataset. The results show TMAD has good accuracy.

Keywords: Text Mining, IDS, Anomaly Detection, TMAD Model, HTTP.

GJCST-G Classification: C.2.1 C.2.0

Strictly as per the compliance and regulations of:

A Text Mining-based Anomaly Detection Model in Network Security

Mohsen Kakavand [α], Norwati Mustapha [σ], Aida Mustapha [ρ] & Mohd Taufik Abdullah [ω]

Abstract- Anomaly detection systems are extensively used security tools to detect cyber-threats and attack activities in computer systems and networks. In this paper, we present Text Mining-Based Anomaly Detection (TMAD) model. We discuss n-gram text categorization and focus our attention on a main contribution of method TF-IDF (Term Frequency, Inverse Document Frequency), which enhance the performance commonly term weighting schemes are used, where the weights reflect the importance of a word in a specific document of the considered collection. Mahalanobis Distances Map (MDM) and Support Vector Machine (SVM) are used to discover hidden correlations between the features and among the packet payloads. Experiments have been accomplished to estimate the performance of TMAD against ISCX dataset 2012 intrusion detection evaluation dataset. The results show TMAD has good accuracy.

Keywords: Text Mining, IDS, Anomaly Detection, TMAD Model, HTTP .

I. Introduction

Changes in network security can be seen as a reliable estimation of transforming trends in the computer science. Information security is a significant issue in present and future life. It also seems to become more important with the development of cyber attacks against social media and mobile devices.

Robertson et al. (2006) referred to several web applications written by individuals with limited knowledge on security. According to CERT/CC, the number of cyber-attacks has increased from 1998 to 2002. Although attacks were relatively few in the early 1990s, a major increase has been reported since 2000 with about 25,000 attacks in 2000 (Malek & Harmantzis, 2004). According to the Common Vulnerabilities and Exposures list (CVE) (Christey & Martin, 2007) and a recent survey on security threats dealing with security risks in digital network world, susceptibility of web application was 25% of the total security issues (Malek & Harmantzis, 2004).

In 1980 James Anderson introduced intrusion detection systems (IDS) (Anderson, 1980) as a counter-action to the dramatic increase of hackers' attacks. There are two kinds IDS (Khalilian, Mustapha, Sulaiman, & Mamat, 2011): misuse detection (MD) and anomaly detection (AD). The latter type generates a model of normal behavior, and removes skeptical behavior or any abnormality from the normal behavior. Anomaly detection is able to identify new attacks, but its main weakness is its vulnerability to false positive alarms. Misuse detection system or signature-based system utilizes knowledge to directly identify the effects of intrusion with high detection precision but is unable to detect new threats and attacks.

Some intrusions use the susceptibilities of a protocol; other attacks attempt to examine a site by probing and scanning. The attacks can be identified by analyzing the network packet headers, or controlling the network traffic connection affairs and session behavior. Patterns in the header fields are such as protocol ID (TCP, UDP, ICMP), Quality-of Service (QoS) flags, port number, and particular values at typical header fields, such as checksums, options, and time to live (TTL). Furthermore, attacks, such as worms, include the delivery of anomaly payload to a susceptible application or service. These attacks might be identified by checking the packet payload. Moreover, examples of payload patterns involve strings such as "GET" and "POST" in the payloads of the HTTP GET and POST request packets. Figure 1 shows the structure of the HTTP GET-request packet, involving the "GET" subsequence pattern.

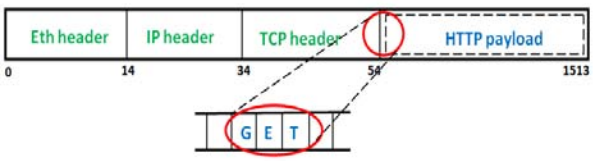


Figure 1: A HTTP GET-request packet

The present study raised the question that, how structural patterns can be identified and characterized? Thus, how can the packets matching those patterns be efficiently recognized? This paper presents a payload anomaly detection model, known as Text Mining-based Anomaly Detection (TMAD) based on data mining / machine leaning techniques.

This paper is organized as follows. In the second section, we review the related works and in the third section, we introduce text mining-based anomaly detection. In the fourth section, we present our system overview with its various subsections. Fifth section discusses the evaluation of TMAD model. Finally, we concluded the paper in the last section.

Author α σ : Faculty of Computer Science and Information Technology, University Putra Malaysia. e-mails : Kakavandirit@gmail.com, {norwati, aida_m, mtaufik} @upm.edu.my.

II. Related Work

Many studies have examined anomaly detection in network traffic. The major obstacles to its practicality are high false-positive rates and lack of clarity and transparency in the detection procedure. The previous methods did not represent adequate precision of false-positive rates. They also did not present diagnostic information to assist forensic analysis.

Some previous methods considered packet header information or statistical properties of sets of packets and connections. Packet header anomaly detection systems such as SPADE (Staniford, Hoagland, & Mcalerney, 2002), PHAD (Mahoney & Chan, 2001), Zhao et al., (2009) (Zhao, Huang, Tian, & Zhao, 2009), Guennoun et al. (2008) (Guennoun, Lbekkouri, & El-khatib, 2008) and Elbasiony et al. (2013) (Elbasiony, Sallam, Eltobely, & Fahmy, 2013) employing statistical approaches for anomaly network traffic and generate alarms when a huge deviation from the normal profile is observed. Feature selection from the packet headers has mostly been used in the mentioned systems. Table 1 summarizes the review on the packet header methods.

Analysis on packet header information typically reduces the data preprocessing necessities. Headers mainly constitute only a tiny part of the whole network data. Thus, processing needs less sources such as storage, memory and CPU. Additionally, features from packet header work quickly, with relatively low memory overheads and computation. Furthermore, they hinder some legal and privacy issues in regard to network packet analysis. Due to these benefits, several studies have utilized packet header as the major features in the intrusion detection systems. However, every request header feature set has its own features.

Therefore, they cannot be used to directly identify attacks bounded for application layer due to the attack bytes are often placed in the request body (Davis & Clark, 2011).

Table 1: Packet header approaches

Authors	Data input	Data preprocessing	Main algorithm	Method	Detection
SPADE, (Staniford et al.,2002)	Packet headers	Preprocessing hold packets with high anomaly score. Score is inverse of probability of packet event.	Entropy, mutual information, or Bayes network.	MD/AD	Probes
PHAD,(Mahoney and chan, 2001)	Ethernet, IP, TCP headers	Models each packet header using clustering	Univariate anomaly detection	AD	Probe, DoS
(Guennoun et al., 2008)	802.11 frame headers	Apply feature construction for 3 higher level features. Feature selection is used to find optimal subset.	K-means Classifier used to detect attacks	AD	Wireless network attacks.
(Zhao et al., 2009)	TCP sessions	Create separate dataset for each application protocol. Quantization of TCP flags within each session.	HMM for HTTP, FTP and SSH to model TCP state transitions.	AD	FTP anomalies
(Elbasiony et al., 2013)	Packet headers	Feature importance values calculated by the random forests algorithm are used in the misuse detection.	Random forests, k- means clustering, weighted k-means	MD/AD	DOS, R2L, U2R, probing

Currently, approaches such as (Estévez-Tapiador, García-Teodoro, & Díaz-Verdejo, 2004), PAYL (Wang & Stolfo, 2004), (Kruegel, Vigna, & Robertson, 2005), McPAD (Perdisci, Ariu, Fogla, Giacinto, & Lee, 2009), SensorWebIDS (Ezeife, Dong, & Aggarwal, 2008) and FARM (Chan, Lee, & Heng, 2013), were suggested for the analysis of packet payloads. These approaches were performed by defining features over payloads, and extracting models of normality based on these features. Packets being not fit into these models are anomalous and trigger alarms. These methods use fairly simple features computed over payload bytes.

Table 2 summarizes the reviewed on packet payload approaches. The table classifies payload anomaly detections works based on the kinds of data preprocessing, major algorithms and detection of various attack categories such as DoS, Buffer overflow, R2L, XML DoS, U2R, etc.

Payload data analysis seems to be more expensive than packet header data analysis as it needs deeper packet inspection, more computation and obfuscation analysis approaches. Due to payload data analysis is complicated, most studies consider tiny subsets of the payload data or only the client-side sections of web content (Davis & Clark, 2011).

This paper aims to examine language models derived from packet payload traffics. Additionally, this study attempts to develop a supervised and

unsupervised learning algorithm that can be directly applied to extracted feature vectors. Thus, the present paper focuses on a major contribution of method TF-IDF (Term frequency, inverse document frequency) for improving an effective computation of measures between text categorization (n-grams).

Table 2. Packet payload approaches

Authors	Data input	Data preprocessing	Main algorithm	Method	Detection
(Tapiador et al., 2004)	Payload	Statistical analysis payload length and mean probability density and standard deviation	Markov chains	AD	HTTP attacks
PAYL(Wag et al., 2004)	Payload	1-g used to compute byte-frequency distribution models for each network destination	Mahalanobis Distance Map (MDM)	AD	Worms, Probe, DoS, R2L, U2R
(Kruegel,et al. 2005)	Web Requests	Construct content-based features from user supplied parameters in URL	Models of normal usage created for each web app. Compare requests to models.	AD/MD	Buffer overflow, Directory traversal, XSS, input validation, Code red
McPAD, (Perdisci, el al., 2009)	Payload	2v-grams extracted from payload. Feature clustering used to reduce dimensionality	One-Class SVM	AD	Shellcode attacks to web servers
SensorWeb IDS (Ezeife, el al., 2008)	Web Requests	network sensor for extracting parameters and the log digger for extracting parameters from web log files	Association Rule Mining (ARM)	AD/MD	XSS, SQL injection, DoS, buffer overflow, cookie Poison
FARM (Chan, el al.,2013)	Payload	Validating User ID, password, service request's input values, input size and SOAP size to from associative patterns and then matching these patterns with interesting rules	Fuzzy Association Rule Mining (FARM)	AD	SQLinjection, XML injection, XML content , SOAP oversized payload, coercive parsing, XML DoS

III. Text Mining-Based Anomaly Detection

Data mining seeks for patterns in data. Similarly, text mining seeks for patterns in text: Analyzing text involves extracting information useful for special goals(Ian H. Witten, Eibe Frank, 2011). Text mining seeks for patterns in natural language texts that are unstructured. Generally, text mining is influential in environments where huge numbers of text documents are managed. Knowledge discovery from text (KDT) considers the machine supported analysis of text. KDT employs methods from information retrieval, information extraction and natural language processing (NLP). Then, it links these three to the algorithms and approaches of Knowledge discovery from data (KDD), data mining, machine learning and statistics (Hotho, Andreas, Paaß, & Augustin, 2005). Text mining is a popular area in anomaly detection which commonly involves the tasks such as clustering, classification, semi-supervised cluster, and so on.

Classification-Based Anomaly Detection analyzes a set of data and generates a set of grouping rules that can categorize future data or predict future data trends called supervised learning. There are different sorts of classification approach such as decision tree induction, Bayesian networks, k-nearest neighbor classifier. In this case, main idea is build a classification model for normal and anomalous events based on labeled training data with require knowledge of both normal and anomaly (attacks) class. The learned model is then applied on the test dataset in order to classify unlabeled records into normal and anomalous records in order to classify each new unseen event (Amer, Goldstein, & Abdennadher, 2013).

Clustering-Based Anomaly Detection is second learning approach called unsupervised learning. Here, the data have no labeling information. Furthermore, no separation into training and testing phase is given. Unsupervised learning algorithms assume that only a small fraction of the data is anomaly and that the attacks exhibit a significantly different behavior than the normal records.

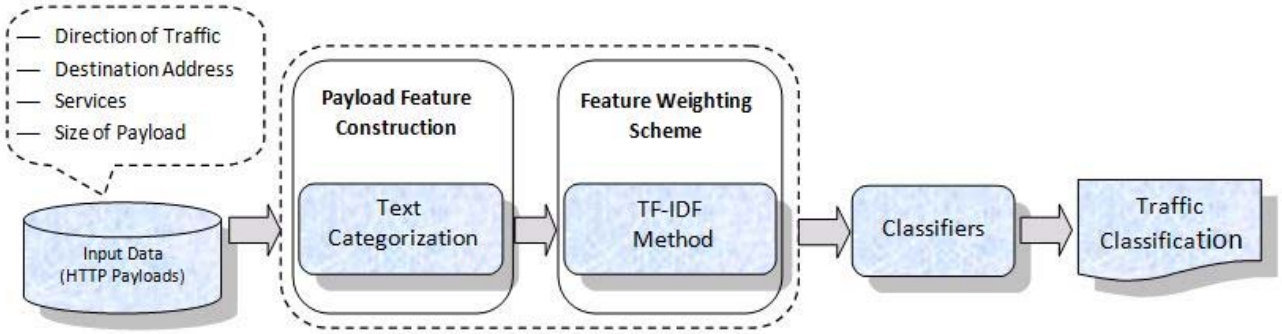


Figure 2 : Overview of text mining based-anomaly detection model

In many practical application domains, the unsupervised learning approach is particularly suited when no labeling information is available. Moreover, in some applications the nature of the anomalous records is constantly changing. Thus, obtaining a training dataset that accurately describe anomaly is almost impossible. On the other hand, unsupervised anomaly detection is the most difficult setup since there is no decision boundary to learn and the decision is only based on intrinsic information of the dataset (Amer et al., 2013).

Semi-Supervised Cluster Analysis, in contrast with classification, is without direction from users. Thus, it may not produce highly valuable clusters. The quality of unsupervised can be highly developed through some weak forms of supervision. Such a clustering is called semi-supervised that is based on user's feedback or guidance constraints is called. (Jiawei Han and Micheline Kamber, 2011). In semi-supervised anomaly detection method, the algorithm models are the only normal records. Records that do not fit into this model are called outliers in the testing stage. Advantages of this semi-supervised anomaly detection can be easily understood of Models as well as normal behavior can be accurately learned but possible high false alarm rate - previously unseen (yet legitimate) data records may be recognized as anomalies (Amer et al., 2013).

However, TMAD model uses a machine learning method and statistical anomaly detection approach to determine an anomaly behavior in the network.

IV. System Overview

This section, represents a comprehensive introduction about the TMAD applying text mining techniques into payload-based anomaly detection. Additionally, the most important contribution of TMAD is the combination of TF-IDF approach and payload-based anomaly detection systems, which have not been investigated in previous studies. This intrusion detection system uses statistical analysis and machine learning algorithms, respectively and then comparison among supervised and unsupervised classifiers. Characters come into network traffic will be analyzed by Mahalanobis Distances Map (MDM) and Support Vector Machine (SVM) to recognize abnormal traffic data from normal ones. Figure 2 shows the overview of TMAD model process.

Text categorization: the text categorization functionality was primarily used. n-Gram text categorization (Wang & Stolfo, 2004), (Banchs, 2013) is in charge of feature construction and request feature analysis. It extracts raw payload features using n-gram (n=1) text categorization method from packet payload and transform observations into a series of feature vectors. Each payload is represented by a feature vector in an ASCII character (256-dimensional).

TF-IDF method: This study investigates the geometrical framework of language modeling. Furthermore, the vector space model and term frequency inverse document frequency (TF-IDF) weighting scheme are examined (Banchs, 2013). Term weighting schemes are often employed to develop the performance. In these schemes, the weights show the significance of a word in a particular document of the selected collection. Huge weights are appointed to terms often used in relevant documents but seldom in the whole document collection (Hotho et al., 2005). Thus, the data resources are processed and the vector space model is set up in order to represent a convenient data structure for text classification. This method is employed to explore similarity between the normal behaviors with the novel input traffic data. The vector space model presents documents as vectors in m-dimensional space, i.e. each document d is explained by a numerical feature vector $W(d) = (x(d, t_1), \ldots, x(d, t_m))$. Therefore, a weight for$W(d, t)$ a term t in document d is computed by term frequency tf (d,t) time inverse document frequency idf (t), describing the term specificity within the document collection. In addition to term frequency and inverse document frequency — defined as (1), a length normalization factor is employed to guarantee that all documents have equal chances of retrieving independent of their lengths (2). Where N is the size of the document collection D and nt is the number of documents in D containing term t.

$$idf_{=}(t) := log(N/n_t) \quad (1)$$

$$W(d,t) = \frac{tf\ (d,t) log\ (N/n_t)}{\sqrt{\sum_{j=1}^{m} idf\ (\mathrm{d,t_j})^2\ (log\ (N/n_{t_j}))^2}}, \quad (2)$$

In the following, the application of geometrical framework model in payload-based anomaly detection is explained.

Classifiers: two distinctive types of algorithms such as Mahalanobis Distances Map (MDM) (Wang & Stolfo, 2004) and Support Vector Machine (SVM) (Scholkopf et al., 1996) are used. Mahalanobis Distances Map (MDM), some factors such as mean value and standard deviation are applied to each byte's frequency. For a payload model, the feature vector that is a set of relative frequencies is the occurrences of each ASCII character to the total number of characters that appear in the payload. Generally, each feature vector can be presented as (3). Then, the mean value and standard deviation of each byte's frequency are computed and explained as (4), (5), (6) and (7) respectively. The mean value and standard deviation vectors, and , are stored in a model M.

$$X = [x_0\ x_1\ \ldots\ x_{255}] \quad (3)$$

$$\bar{X}\ = [\bar{x}_0\ \bar{x}_1\ \ldots\ \bar{x}_{255}] \quad (4)$$

$$\bar{\sigma}\ = [\bar{\sigma}_0\ \bar{\sigma}_1\ \ldots\ \bar{\sigma}_{255}] \quad (5)$$

$$\bar{x}_i = \frac{1}{n} \sum_{k=1}^{n} x_{i,k} \quad (0 \le i \le 255) \quad (6)$$

$$\bar{\sigma}_i = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x}_i)^2} \quad (0 \le i \le 255) \quad (7)$$

The model considers the correlations among various features (256 ASCII characters). Therefore, for each network packet, a feature vector is defined by (3). Here, there are the average value of features in the 1-gram model (8) and the covariance value of each feature (9). To investigate the association among the characters, where μ is the average frequency of each ASCII character presented in the payload, Σi is the covariance value of each feature. Next, the classifiers between two characters (10) are presented:

$$\mu = \frac{1}{256} \sum_{i=0}^{255} x_i \quad (8)$$

$$\sum \boldsymbol{i}\ = (x_i - \mu)\ (x_i - \mu)'\ (0 \le i \le 255) \quad (9)$$

$$\boldsymbol{d}_{(i,j)} = \frac{(x_i - x_j)\ (x_i - x_j)'}{\sum \boldsymbol{i} + \sum \boldsymbol{j}} \quad (0 \le i \le 255) \quad (10)$$

MDM is a statistical method or a model-based method that is created for the data. The objects of the study are evaluated with respect to how well they fit into the model. According to the above evaluation, the MDM of a network packet is made as matrix D. Then, the Mahalanobis distance between two distributions of D and the model *M* is evaluated. Then, the weight w is calculated using (11), if the weight is larger than a threshold, the input packet is considered as an intrusion.

$$W = \sum_{i,j}^{255,255} \frac{(d_{ob\ (i,j)} - \bar{d}_{nor\ (i,j)})2}{\sigma^2_{nor\ (i,j)}} \quad (11)$$

In the following, the Support Vector Machine (SVM) algorithm is chosen. This algorithm is originally proposed by Scholkopf et al. in (Scholkopf et al., 1996). SVM have been shown to achieve good performance in text mining classification problems (Sebastiani, 2002). It is also known as a supervised classification algorithm that is able to process feature vectors of high dimensions for providing a fast and influential approach for learning text classifiers (LEOPOLD & KINDERMANN, 2002). Typically, document d is presented by a – possibly weighted – vector $(t_{d1}, \ldots t_{dN})$ of the counts of its featur. A SVM can only separate two groups — a positive group L1 (shown by y = +1) and a negative group L2 (shown by y = -1). In input vectors, a hyperplane might be defined by setting y = 0 as follows:

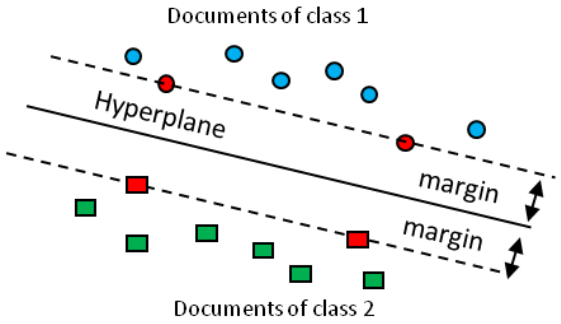


Figure 3 : Decision boundary and margin of SVM

The SVM algorithm identifies a hyperplane placed between the positive and negative examples of the training set. Figure 3 shows the parameters bj are adapted in a manner that the distance ε– called margin – between the hyperplane and the closest positive and negative example packet payload is maximized. This considers a limited quadratic optimization issue which can be resolved for a huge number of input vectors. The documents that have distanceε from the hyperplane are known as support vectors and identify the actual location of the hyperplane (12).

$$y = \int (\overrightarrow{t_d}) = b_0 + \sum_{j=1}^{N} b_j t_{dj} \quad (12)$$

Typically only a tiny fraction of payloads are considered as support vectors. A new document with term vector $\overrightarrow{t_d}$ is classified in L1 if the value $\int (\overrightarrow{t_d}) > 0$ and into L2 otherwise. If the packet payload vectors of the two classes are not linearly separable, a hyperplane is selected. Classifier approaches identify patterns of packet payloads in network traffic data. These are undertaken in extracting the hidden correlations

between features and the correlations among network packet payloads.

V. Experimental Results

This section represents results of the extensive experiments performed. This study examined TMAD model on the ISCX dataset 2012 (Shiravi, Shiravi, Tavallaee, & Ghorbani, 2012), reflecting current trends traffic patterns and intrusions. This is in contrast to static datasets that are widely used today but are outdated, unmodifiable, inextensible, and irreproducible. ISCX dataset 2012 is considered as a new standard data set for evaluation of intrusion detection systems.

a) ISCX Dataset 2012

To evaluate TMAD model, ISCX dataset 2012 (Shiravi et al., 2012) were collected under the sponsorship of Information Security Centre of Excellence (ISCX). All the network traffic of the data set was included in both normal network traffic and attack traffic for system evaluation of the proposed approach for text mining based-anomaly detection. ISCX dataset 2012 contains categories of attacks including scan, DoS, R2L, U2R and DDoS.

The entire ISCX labeled dataset comprises nearly 1512000 packets and covered seven days of network activity. However, the ready-made training and testing dataset is not available. Thus, ISCX HTTP/GET traffic was randomly divided into two parts: a training set made of approximately 80% of the HTTP/GET traffic and a testing set made of the remaining 20% of the traffic. The present study focuses on the anomalous coming through inbound HTTP requests. HTTP-based attacks are mainly from the HTTP GET request at the server side. Request bodies carry data to the web server but sometimes the request message has no body, because no request message data is needed to give GET a simple document from a server (Davide & Brian, n.d.). Therefore, we removed all HTTP request packets begin without request message.

b) Analysis And Result

We directed experiments for our TMAD model with the extracted data from the ISCX dataset 2012. In the first part of our experiments, we present the model generation for normal HTTP traffic. Afterwards, we evaluate the accuracy of our TMAD model in detecting various attacks coming through HTTP services including scan, DoS, R2L, U2R and DDoS. We trained the TMAD model on the ISCX dataset 2012, and then evaluate the model on test dataset, which contains different attacks. For port 80, the attacks are often malformed HTTP requests and are very different from normal requests.

For our research, we have plan on the anomalous incoming through inbound HTTP requests. HTTP-based attacks are mostly from the HTTP GET request at the server side. Figure 4 shows histogram of HTTP request distributions.

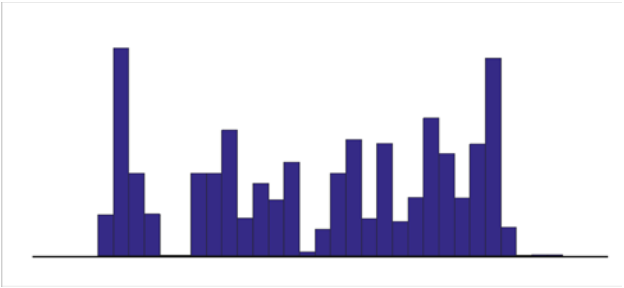


Figure 4 : Histogram of HTTP request distributions

Figure 5 provides an example displaying the variability of the frequency distributions from port 80. The plot represents the characteristic profile for that port 80 and flow direction (inbound) full length payloads.

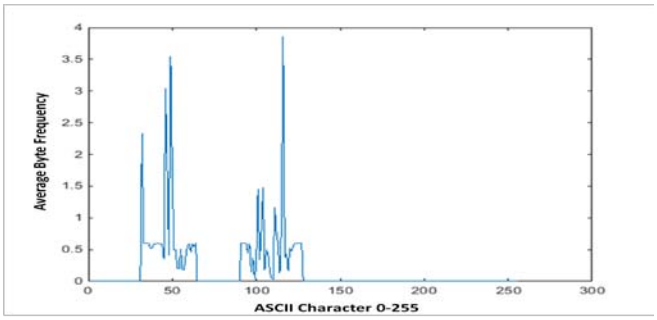


Figure 5 : Example byte distribution for port 80

As illustrated in figure 6, (a) and (b) indicate the anomaly free and anomaly character relative frequencies. We can see the character relative frequencies of anomaly packet payloads are very different from the normal packet payloads, which can distinguish anomalous from normal packet payloads.

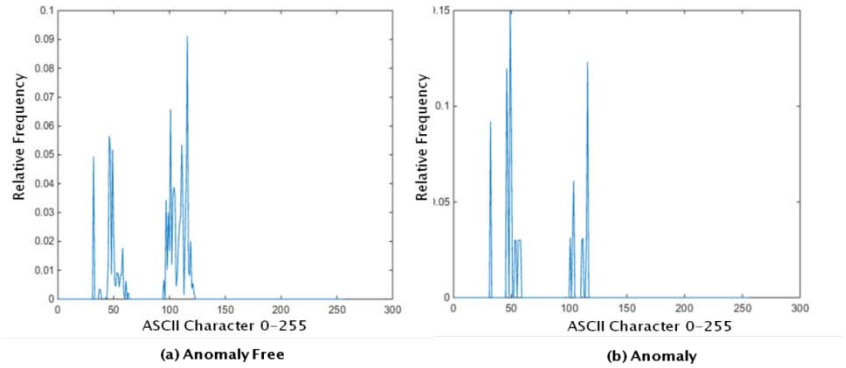


Figure 6 : Relative Frequencies of Characters of Normal and Anomaly

Concurrently, the experimental results illustrated the acceptable performance of the TMAD model in detecting HTTP anomaly on web services. That is the knowledge from the TF-IDF method explaining the correlation among 256 ASCII characters. The results showed that the Text Mining-Based Anomaly Detection (TMAD) is able to detect new attacks with different detection rate and false positive rate in MDM and SVM algorithms. We also used Receiver Operating Characteristic (ROC) curve method to compare the performance of our model with MDM and SVM. The ROC curve showed incorrectly flagging non-attack requests as an attack (false positives) and detection rate attack. TP and FP rate are shown in figure 7 in ROC curve.

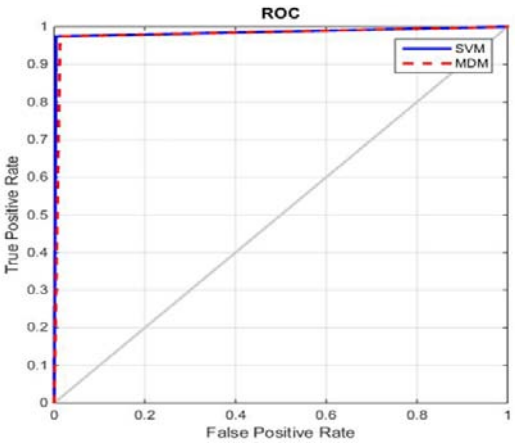


Figure 7 : ROC Curve for the Accuracy of the TMAD model

The results obtained for the model are very encouraging. TMAD has a detection rate around 97.44% and 1.3% false positive rates for MDM and detection rate around 97.45% with 0.4% false positive rates for SVM. Performance obtained by TMAD model in training and test data is presented in Table 3.

Table 3 : MDM and SVM using Training and Testing Dataset

Algorithm	Training Data		Testing Data	
	Detection Rate	False Positive	Detection Rate	False Positive
MDM	93.46%	3.3%	97.44%	1.3%
SVM	97.61%	1.6%	97.45%	0.4%

VI. Conclusions

This study presented TMAD (Text Mining-Based Anomaly Detection), a new approach to detect HTTP attacks in the network traffic. TMAD is an anomaly detector based upon text categorization and TF-IDF method. The use of TF-IDF improves the performance usually term weighting schemes are used, where the weights reflect the importance of a word in a specific document of the considered collection.

The experimental result indicated that the method is effective at detection rate, but the notable detection accuracy suffered from high level of false positive rate for ISCX dataset 2012 (Shiravi et al., 2012) collected under the sponsorship of Information Security Centre of Excellence (ISCX). For port 80, it achieved almost 97.44% detection rate with around 1.3% false positive rate in unsupervised learning method, and 97.45% detection rate with around 0.4% false positive rate in supervised learning. In future research, we intend to reduce the dimensionality of feature space and false positive rate by applying data-mining preprocessing techniques to ISCX dataset 2012.

In our future work we aim to evaluate the performance of TMAD model on 1999 DARPA/MIT Lincoln Laboratory, which produced the most prominent datasets for testing IDS. Moreover, we will try to reduce of false positive rate and improve detection rate.

VII. Acknowledgement

This research work supported by the Fundamental Research Grant Scheme (FRGS) under Ministry o Higher Education, Malaysia and it is gratefully appreciated.

References Références Referencias

1. Amer, M., Goldstein, M., & Abdennadher, S. (2013). Enhancing one-class support vector machines for unsupervised anomaly detection. Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description- ODD '13, 8–15. doi:10.1145 /2500853.2500857
2. Anderson, J. P. (1980). Computer Security Threat Monitoring And Surveillance (pp. 1–56). Washington.
3. Banchs, R. E. (2013). Text Mining with MATLAB® (pp. 1–347). New York, NY: Springer New York. doi:10.1007/978-1-4614-4151-9
4. Chan, G.-Y., Lee, C.-S., & Heng, S.-H. (2013). Discovering fuzzy association rule patterns and increasing sensitivity analysis of XML-related attacks. Elsevier Science, Journal of Network and Computer Applications, 3
5. Christey, A. S., & Martin, R. A. (2007). Vulnerability Type Distributions in CVE. MITRE, Common Weakness Enumeration (CWETM). Retrieved May 22, 2007, from http://cwe.mitre.org/documents/vuln-trends/index.html
6. Davide, G., & Brian, T. (n.d.). HTTP The Definitive Guide (2nd ed., p. 617). United States of America: Publisher: O'Reilly Media.
7. Davis, J. J., & Clark, A. J. (2011). Data preprocessing for anomaly based network intrusion detection: A review. Computers & Security, 30(6-7),353–375. doi:10.1016/j.cose.2011.05.008
8. Elbasiony, R. M., Sallam, E. a., Eltobely, T. E., & Fahmy, M. M. (2013). A hybrid network intrusion detection framework based on random forests and weighted k-means. Ain Shams Engineering Journal, 4(4), 753–762. doi:10.1016/j.asej.2013.01.003
9. Estévez-Tapiador, J. M., García-Teodoro, P., & Díaz-Verdejo, J. E. (2004). Measuring normality in HTTP traffic for anomaly-based intrusion detection. Elsevier, Computer Networks, 45(2), 175–193. doi:10.1016/j.comnet.2003.12.016
10. Ezeife, C. I., Dong, J., & Aggarwal, A. K. (2008). SensorWebIDS: a web mining intrusion detection system. Emerald, International Journal of Web Information Systems, 4(1), 97–120. doi:10.1108/17440080810865648
11. Guennoun, M., Lbekkouri, A., & El-khatib, K. (2008). Selecting the Best Set of Features for Efficient Intrusion Detection in 802 . 11 Networks. In in:Information and communication technologies: from theory to applications, ICTTA 2008. 3rd International Conference. (pp. 1–4).
12. Hotho, A., Andreas, N., Paaß, G., & Augustin, S. (2005). A Brief Survey of Text Mining. LDV Forum - GLDV Journal for Computational Linguistics and Language Technology, 1–37.
13. Ian H. Witten, Eibe Frank, M. A. H. (2011). Data Mining Practical Machine Learning Tools and Techniques (Edition, T., Vol. 40, pp. 1–629). Elsevier. doi:10.1002/15213773(20010316)40:<923 ::AID-ANIE9823>3.3.CO;2-C.
14. Jiawei Han and Micheline Kamber. (2011). Data Mining: Concepts and Techniques (Third Edit.). San Francisco, CA, USA: Morgan Kaufmann.
15. Khalilian, M., Mustapha, N., Sulaiman, N., & Mamat, A. (2011). Intrusion Detection System with Data Mining Approach: A Review. Global Journal Of Computer Science & Technology, 11(5).
16. Kruegel, C., Vigna, G., & Robertson, W. (2005). A multi-model approach to the detection of web-based attacks. Elsevier Science, Computer Networks, 48(5),717–738. doi:10.1016/j.comnet.-20 05.01.009
17. LEOPOLD, E., & KINDERMANN, J. (2002). Text Categorization with Support Vector Machines .How to Represent Texts in Input Space ? Kluwer Academic Publishers Machine Learning, (22), 423–444.
18. Mahoney, M. V, & Chan, P. K. (2001). PHAD : Packet Header Anomaly Detection for Identifying Hostile Network Traffic. Florida Institute of Technology Technical Report, (1998), 1–17.
19. Malek, M. ;, & Harmantzis, F. (2004). Data mining techniques for security of web services. In Proc. International Conference on E-Business and Telecommunication Networks (ICETE) (Vol. 1–14, pp. 1–14). Retrieved from http://www.stevens-tech.edu/perfectnet/publications/Papers/ICETE2004 -paper.pdf
20. Perdisci, R., Ariu, D., Fogla, P., Giacinto, G., & Lee, W. (2009). McPAD : A Multiple Classifier System for Accurate Payload-based Anomaly Detection. Elsevier Science, Computer Networks, 5(6), 864–881.
21. Robertson, W., Vigna, G., Kruegel, C., & Kemmerer, R. A. (2006). Using Generalization and Characterization Techniques in the Anomaly-based Detection of Web Attacks. In: Proceedings of the Network and Distributed System Security Symposium (NDSS), 14. Retrieved from http://iseclab.org/papers/webfuzzing.pdf
22. Scholkopf, B., Sung, K., Burges, C., Girosi, F., Niyogi, P., Poggio, T., & Vapnik, V. (1996). Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers. IEEE Transactions on Signal Processing, (1996-12-01). doi:10.1109/78.650102
23. Sebastiani, F. (2002). Machine learning in automated text categorization. ACM Computing Surveys, 34(1), 1–47. doi:10.1145/505282.505283
24. Shiravi, A., Shiravi, H., Tavallaee, M., & Ghorbani, A. a. (2012). Toward developing a systematic

approach to generate benchmark datasets for intrusion detection. Computers & Security, Elsevier, 31(3), 357–374. doi:10.1016/j.cose.2011.12.012

25. Staniford, S., Hoagland, J. A., & Mcalerney, J. M. (2002). Practical automated detection of stealthy portscans. Journal of Computer Security, 10, 105–136.
26. Wang, K., & Stolfo, S. J. (2004). Anomalous Payload-based Network Intrusion Detection. Springer Berlin Heidelberg,7th International Symposium, RAID, Sophia Antipolis, France, September 15 - 17, Proceedings, pp 203–222.
27. Zhao, J., Huang, H., Tian, S., & Zhao, X. (2009). Applications of HMM in Protocol Anomaly Detection. In 2009 International Joint Conference on Computational Sciences and Optimization (pp. 347–349). Ieee. doi:10.1109/CSO.2009.51

This page is intentionally left blank

Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 14 Issue 5 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

Priority based Congestion Control Mechanism in Multipath Wireless Sensor Network

By Md. Ahsanul Hoque, Nazrul Islam, Sajjad Waheed & Abu Sayed Siddique

Mawlana Bhashani Science and Technology University, Bangladesh

Abstract- Wireless Sensor Network (WSN) is a network composed of distributed autonomous devices using sensors. Sensor nodes send their collected data to a determined node called Sink. The sink processes data and performs appropriate actions. Nodes using routing protocol determine a path for sending data to sink. Congestion occurs when too many sources are sending too much of data for network to handle. Congestion in a wireless sensor network can cause missing packets, long delay, overall channel quality to degrade, leads to buffer drops. Congestion control mechanism has three phases, namely congestion detection, congestion notification and congestion control. In this paper is propose two bit binary notification flag to notify the congested network status for implicit congestion detection. For congested network status, we propose a priority based rate adjustment technique for controlling congestion in link level. Congested packet will be distributed equally to the child node to avoid packet loss and transition delays based on technique.

Keywords: congestion control, multipath, weighted fairness, queue overflow and channel overloading.

GJCST-G Classification: C.2.1 I.2.9

Strictly as per the compliance and regulations of:

Priority based Congestion Control Mechanism in Multipath Wireless Sensor Network

Md. Ahsanul Hoque[α], Nazrul Islam [σ], Sajjad Waheed[ρ] & Abu Sayed Siddique [ω]

Abstract- Wireless Sensor Network (WSN) is a network composed of distributed autonomous devices using sensors. Sensor nodes send their collected data to a determined node called Sink. The sink processes data and performs appropriate actions. Nodes using routing protocol determine a path for sending data to sink. Congestion occurs when too many sources are sending too much of data for network to handle. Congestion in a wireless sensor network can cause missing packets, long delay, overall channel quality to degrade, leads to buffer drops. Congestion control mechanism has three phases, namely congestion detection, congestion notification and congestion control. In this paper is propose two bit binary notification flag to notify the congested network status for implicit congestion detection. For congested network status, we propose a priority based rate adjustment technique for controlling congestion in link level. Congested packet will be distributed equally to the child node to avoid packet loss and transition delays based on technique. Furthermore, this technique allocates the priority of many applications simultaneously running on the sensor nodes, which route is own data as well as the data generated from other sensor nodes. The results show that the proposed technique achieves better normalized throughput and total scheduling rate with the avoiding packet loss and delay.

Keywords: congestion control, multipath, weighted fairness, queue overflow and channel overloading.

I. Introduction

Wireless sensor network typically has little or no infrastructure. It consists of a number of sensor nodes (few tens to thousands) working together to monitor a region to obtain data about the environment. WSN has gained worldwide attention in recent years. These sensor nodes can sense, measure to gather information from the environment. Based on some local decision process, they can transmit the sensed data to the user. A variety of mechanical, thermal, biological, chemical, optical, and magnetic sensors may be attached to the sensor node to measure properties of the environment. WSN consists of spatially distributed autonomous sensor nodes to cooperatively monitor physical or environmental conditions.

The sensor nodes of a WSN sense the physical phenomena and transmit the information to base stations. When an event occurs, the load becomes heavy and the data traffic also increases. This might lead to congestion.

There are mainly two causes for congestion in WSN. The first case is node level congestion, which occurs when the packet-arrival rate exceeds the packet-service rate. This is more likely to occur at sensor nodes close to the sink, as they usually carry more combined upstream traffic. The second case is link level congestion, which occurs due to contention, interference, and bit-error rate. Congestion control mechanism has three phases: congestion detection, congestion notification and congestion control with rate adjustment technique.

In recent years, lots of work going on in congestion control for wireless sensor network. Most of the work deals with the priority based rate adjustment algorithm for different types of application for heterogeneous traffic. Congestion Control and Fairness for Many-to-one Routing in Sensor Networks [1] proposes a distributed and scalable algorithm that eliminates congestion within a sensor network and that ensures the fair delivery of packets to a central node or base station. Priority Based Congestion Control for Heterogeneous Traffic in Multipath Wireless Sensor Networks [2] proposes a priority based congestion control for heterogeneous traffic in multi path wireless sensor network. The proposed protocol allocates bandwidth proportional to the priority of many applications simultaneously running on the sensor nodes. Congestion Detection and Avoidance (CODA) [4] uses buffer occupancy and the channel load for measuring congestion level. It handles both transient and persistent congestion. For transient congestion, the node sends explicit backpressure messages to its neighbors were as for persistent congestion; it needs explicit ACK from the sink. It uses three mechanisms as follows: 1) Receiver based congestion detection, 2) open loop, hop-by hop backpressure, 3) closed loop multi source regulation. PCCP (Priority Based Congestion Control Protocol) [7] is a node priority based congestion control protocol. It defines priority from the nodes point of view instead of the traffic flows point of view.

For implicit congestion detection, we proposed two bit binary flags such as 00, 01, 10 or 11 for congestion notification. Based on congested or heavily loaded network status, we propose different activities

Author α σ ρ ω: Department of Information and Communication Technology Mawlana Bhashani Science and Technology University, Tangail-1902, Bangladesh, e-mails : ahsan.ict156@gmail.com, nazrul.islam@mbstu.ac.bd, sajad302@yahoo.com, hemel.ict.mbstu@gmail.com

such as pass packet with delay factor or pass packets with rate adjustment technique. In rate adjustment technique, here also calculated with a delay factor. The rate adjustment technique also distributes congested packet, equally to all child nodes to avoid packet loss.

The structure of this paper is as follows. Section II represents the research methodology of the study. Section III represents details about the technical implementation. Section IV provides result and discussion on the basis of technical implementation of the research. This section also shows a comparison between simple fairness and weighted fairness. Finally, section V summarized a set of conclusion and through an outlook of the future work.

II. Research Methodology

A literature review to find about wireless sensor network and also find about congestion in it. Congestion in wireless sensor network causes overall channel quality to degrade and loss rates rise, leads to buffer drops, packet loss and increased delays. Most of the work in recent years has been done in congestion control to avoid packet loss or to minimize packet loss.

In order to find a congestion control technique which avoid packet loss and capable of reducing delay. In these circumstances we discover a system architecture which contains congestion detection unit, congestion notification unit and congestion controlling unit. The analytical data provide different network status from which we can make a decision whether the congestion is occurring or not.

Furthermore, we found some analytical data from the multipath multihop network model. From the basis of these analyses data, a plot represents the comparison of simple fairness, weighted fairness and throughput. Finally, careful study of the plots is expected to provide a fairness measure of the effect of different packets.

III. Technical Implementation

We discuss in this section the detail of our proposed work namely.

a) *System Architecture*

In Figure 1 represents the system architecture of the proposed work. The Congestion Detection Unit (CDU) calculates the packet service ratio [8]. With the help of congestion control Unit each packet, equally distributed to the child node with existing priority allocates the bandwidth to the child nodes according to the source traffic priority and transit traffic priority. The Congestion Notification Unit (CNU) uses an implicit congestion notification by piggybacking the rate information in its packet header. All the child nodes of a parent node overhear the congestion notification information[12].

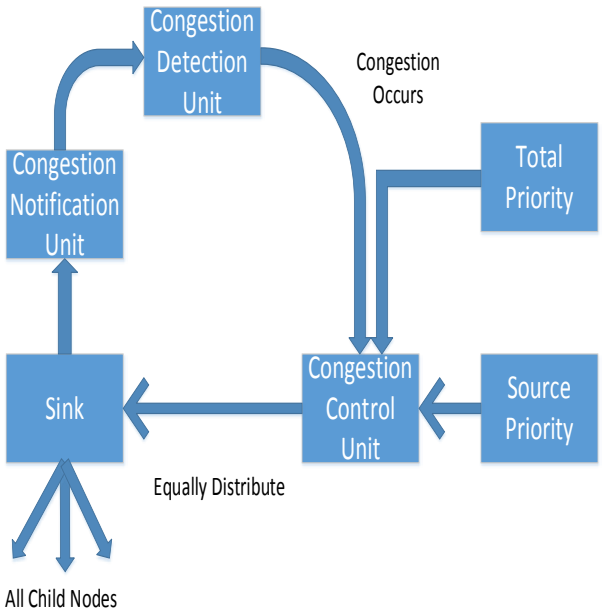


Figure1: Side Channel Power Analysis Attack

b) *Queuing Model*

In Figure 2 shows the queuing model of each sensor node. To differentiate different types of traffic in the heterogeneous network, source sensor node adds a traffic class identifier to identify the traffic class [9]. For each traffic class a separate queue is maintained in the sensor nodes. The classifier classifies the packets based on the traffic class and sends them to the matching unit to set them in a single queue.

The following are the descriptions of the queuing model.

- Source Rate r_s^i: It is the rate at which a sensor node originates data.
- Scheduling rate r_{sch}^i: It is defined as the rate at which the scheduler schedules the packets from the queues. The scheduler of node i, forwards the packet from node i-1 to the next node i+1.

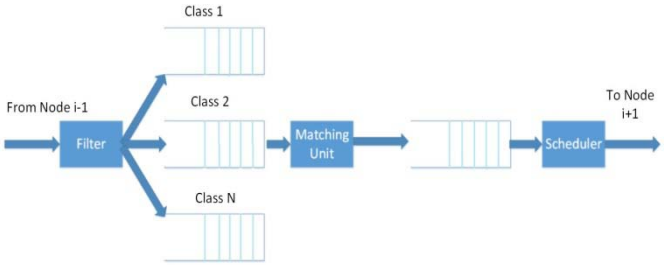


Figure 2 : Queuing Model

The scheduler schedules the packets from queues according to the queue priority [10]. The packets from a higher priority queue will be serviced more than the packets from the lower priority queue. The packets in a particular queue are processed based on its sourced traffic and transit traffic [3]. Transit traffic gives more priority than source traffic, since the transit traffic data have already been traversed several paths, and dropping them would cause more waste of network resources. The classifier differentiates the transit traffic from source traffic by examining the source address in the packet header [11].

Congestion control mechanism containing three phase congestion detection, congestion notification and finally congestion control with rate adjustment technique [5]. Theoretically, we implement our proposed technique that avoids congestion occurrence and minimize packet loss.

i. *Congestion Detection*

For congestion detection we use, congestion notification flag based on different types of data stored in a single queue to pass the packet through the channel to destination. Sensor node may have multiple sensors with different types of characteristics stored in a single queue. Based on queue for different types of data with different characteristics we use, congestion notification flag such as if the queue is lightly loaded than it notify 00 and the packet pass through the channel without any modification, if it is loaded then it notify 01 and the packet also pass through the channel with queue priority. If the queue is heavily loaded, then it notifies 10 and we propose a delay to avoid congestion in our mechanism and if the queue notify 11 means congestion than we control it with our proposed priority based rate adjustment technique.

ii. *Congestion Notification*

Congestion notification is calculated based on queue size. If the queue average less than queue minimum than it notify 00, If the queue average greater than or equal queue minimum or less than queue warning value than it notify 01, If the queue average less than or equal queue maximum or greater than equal warning, it notifies 10, otherwise queue average greater than the queue minimum than it notify 11 means congestion occurred.

Table1: Congestion Notification Flag and Action Status

CN Flag	Network Status	Action
00	Lightly loaded	Pass packet without modification
01	Loaded	Pass packet based on queue priority
10	Heavily loaded	Using delay factor & pass packet
11	Congested	Controlling with rate adjustment technique

iii. *Rate Adjustment Technique*

Rate adjustment is done in the single queue. It ensures that the heterogeneous data from different queue will reach the base queue at the desired rate.

When each node receives the congestion notification information, it adjusts the queue rate accordingly [14].

The following describe the rate adjustment technique:

The packet service ratio r (i) is used to measure the congestion level at each node i. The packet service ratio is calculated as follows:

$$r(i)=rs(i)/rsch(i)$$

Here r(i) denoted as packet service ratio used to measure the congestion level for each node where rs(i) denotes packet service rate of node i and rsch(i) denotes packet scheduling rate of node i.

The average service time can be calculated as follows by using Exponential weighted moving average formula.

$$\overline{T}_s(i)=(1-k)\overline{T}_s(i)+KT_s(i)$$

Where K is a constant value in the range between $0<K<1$ and TS(i) denotes the service time of the current packet in sink node. By using EWMA formula is updated each time a packet is forwarder to the next [6]. The average packet service rate is calculated as the inverse of the average service time.

$$R^i_{serv}=\frac{1}{\overline{T}s(i)}$$

If $r(i)<Qmin$ normal operation. Else if $r(i)>=Qmin$ & $r(i)<Qavr$ prefer queue priority and pass packet. Else if $r(i)>=Qavr$ & $r(i)<Qwarn$ pass packet with delay to avoid congestion. Else $r(i)>=Qmax$ distribute packet to the child node.

Let the traffic source priority SP(i) for parent node i of application j where j as child node and i as sense node. For each node i, total traffic source priority can be calculated as the sum of all application running in it.

$$SP(i)=\sum_{j=1}^{n}\left(SP_i^j\right)$$

Where SPij denotes the traffic source priority of application j. The symbol n denotes the number of application running in sensor node i under multipath routing. The packets of a flow may pass through multiple paths before they arrive at the sink and only a fraction of a flow passes through a particular node or link. The traffic of a particular flow forwarded in one path is defined as a sub flow [15]. Transit traffic priority at sensor node i is used to represent the relative priority of transit traffic from other nodes routed through node i.

Let dlyji for application j in sense node i carries priority that depends on child node that means single node to single node or single node to multi node. If we denote transit priority TP(i) for node i than it would be

$$TP(i)=dly^j_i\times\sum_{j=0}^{j=1}GP(j)$$

$$GP(i)=SP(i)+TP(i)$$

Where GP(i) stands for global priority for node i , Here (dlyji) calculation depends on node number, here we assume that if single node to single node, the value of dlyji must be 1, otherwise it is calculated by dividing delay with the number of parent node connected with [13]. Where SP(i) denotes the source priority for each node carrying with. The process running continuously to pass packet to the sink varied with transit priority and global priority. The packet service ratio reflects the congestion level at each sensor node. When this ratio is equal to 1, the scheduling rate is equal to the service rate. When this ratio is greater than 1, the scheduling rate is less than the packet service rate. In both these cases, there is no congestion. When the packet service ratio is less than 1, the scheduling rate is more than service rate and it causes the queuing up of packets. It indicates congestion. Let threshold value is 0.75. If the packet service ratio is less than the threshold value, it notifies congestion. Then node i will adjust the scheduling rate. If the packet service ratio is greater than 1, indicates that the packet service rate is greater than the scheduling rate. So each node will increase the scheduling rate for parent j to improve link utilization otherwise the packet service rate is equal to the scheduling rate.

$$r_{sch}(i)=GP(i)/(GP(i-1)+GP(i)+GP(i+1))$$
$$Tr_{sch}(i)=r_{sch}(i)/(r_{sch}(i-1)+r_{sch}(i)+r_{sch}(i+1))$$

Scheduling rate can be calculated based on threshold and its directly connected to service ratio to pass packet, and if the service ratio r(i) exceeds the queue maximum value distribute the packet to the child node to avoid packet loss and continue the process to avoid congestion.

IV. Result and Discussion

All the child nodes of the node i overhear the congestion notification information and they control their rate according to it. Rate adjustment is done in a queue. It ensures that the heterogeneous data from different class will reach the base queue at the desired rate. When the queue explodes the congestion notification flag, it adjusts its rate accordingly. Figure 3, shows the multipath multi hop heterogeneous network model considered in the network. In case of multipath routing, each node divides its total traffic into multipath flows and those flows pass through multiple downstream nodes. For simplicity, we assume that each node divides its rate equally among all its parents. We assume that application 1 generates traffic of class 1 and its priority is 1 and for application 2 the priority is 2 and so on. Then the total priority is calculated based on the traffic class of application running on the nodes. The rate allocates to each node is calculated based on the total priority.

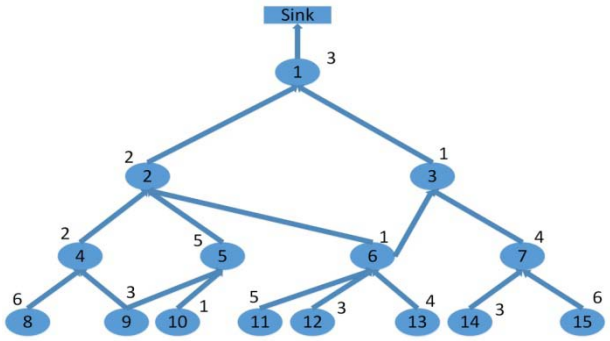


Figure 3 : Multipath multi hop network model

Table 2 : Congestion Notification Flag and Action Status

Node	App1	App2	App3	SP(i)
1	0	0	1	3
2	0	1	0	2
3	1	0	0	1
4	0	1	0	2
5	0	1	1	5
6	1	0	0	1
7	1	0	1	4
8	1	1	1	6
9	0	0	1	3
10	1	0	0	1
11	0	1	1	5
12	1	1	0	3
13	1	0	1	4
14	0	0	1	3
15	1	1	1	6

Table II shows the source traffic priorities of nodes i in heterogeneous environments.

As shown in line 5, node 5 has applications 2 and 3 running on it, which generates heterogeneous data with traffic class priority 2 and 3 respectively. The source traffic priority of node 5 is 0+2+3 = 5.

The source traffic priority of all the other nodes can be calculated by the sum of source traffic priorities of traffic classes of individual applications running in it. Node 2 routes the traffic from 4, 5 and 6. The delay factor must be 1/3. The transit traffic priority of node 2 is calculated from the global priorities of nodes 4, 5 and 6.

$$TP(i) = dly_i^j \times \sum_{j=0}^{j=1} GP(j)$$

$$TP(2) = 1/3 \times 29 = 10$$

Table 3 : Scheduling Rate and The Normalized Throughput of The Sensor Nodes

Node	TP(i)	dly_i^j	GP (i)	$r_{sch}(i)$	$Tr_{sch}(i)$	Normalizd throughput of node i
1	12	1	15	0.555	0.643	0.643
2	10	0.33	12	0.307	0.254	0.254
3	11	0.5	12	0.342	0.344	0.344
4	9	0.5	11	0.343	0.344	0.344
5	4	0.5	9	0.310	0.328	0.328
6	8	0.33	9	0.290	0.295	0.295
7	9	0.5	13	0.382	0.369	0.369
8	6	1	12	0.387	0.362	0.362
9	3	1	6	0.334	0.376	0.376
10	1	1	2	0.112	0.114	0.114
11	5	1	10	0.556	0.60	0.60
12	3	1	6	0.254	0.208	0.208
13	4	1	8	0.445	0.454	0.454
14	3	1	6	0.230	0.177	0.177
15	6	1	12	0.670	0.730	0.730

Table III. shows the scheduling rate and the normalized throughput of the sensor nodes.

Figure 4.compares the normalized throughput obtained from the proposed technique for priority based weighted fairness with the simple fairness case. As shown in Figure 4, in simple fairness when the priorities of all nodes are same, they receive equal throughput. The proposed method provides priority based weighted fairness for all the sensor nodes according to the priority of heterogeneous traffic generated by the applications.

The rate allocation and scheduling are made according to the priority. The node 10 which has higher priority has highest normalized throughput and nodes 15 and 1 got lowest normalized throughput. Thus, the proposed mechanism allocates the bandwidth to each sensor node based on its priority.

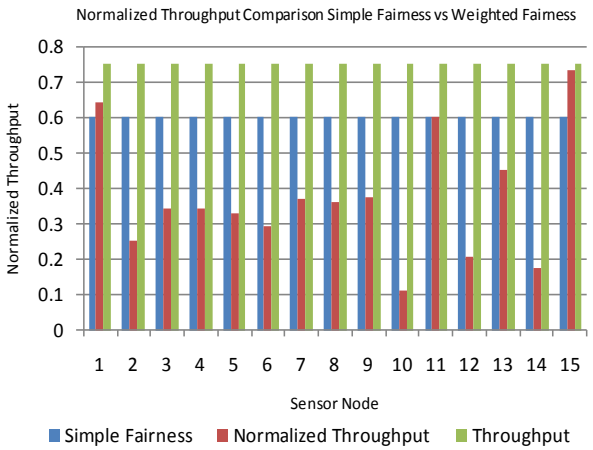


Figure 4 : Normalized Throughput Comparison

V. Conclusions

In this paper is proposed a priority based congestion control mechanism for heterogeneous data for multipath environment. We have calculated source priority depending on running applications of the sensor nodes and also calculates transit priority by multiplying with delay factor and global priority. In the proposed method the queue divides its congested packet to the child nodes equally and then running the controlling mechanism to prevent or avoid congestion. The proposed technique minimizes congestion occurrence through packet drop ratio, delay and normalized throughput.

In future work, a real time application will be made and the performance will be evaluated through congestion avoidance technique. It is anticipated that will be no packet loss or any kind of delay in packet transmission.

References Références Referencias

1. Ee, C. T., &Bajcsy, R. (2004, November). Congestion control and fairness for many-to-one routing in sensor networks. In *Proceedings of the 2nd international conference on Embedded networked sensor systems* (pp. 148-161). ACM.
2. Sridevi, S., Usha, M., &Lithurin, G. P. A. (2012, January). Priority based congestion control for heterogeneous traffic in multipath wireless sensor networks. In *International Conference on Computer Communication and Informatics* (ICCCI), 2012 (pp. 1-5). IEEE.
3. Monowar, M. M., Rahman, M. O., & Hong, C. S. (2008, February). Multipath congestion control for heterogeneous traffic in wireless sensor network. In *10th International Conference on Advanced Communication Technology, 2008.ICACT 2008.*(Vol. 3, pp. 1711-1715).IEEE.
4. Wan, C. Y., Eisenman, S. B., & Campbell, A. T. (2003, November). CODA: congestion detection and avoidance in sensor networks. In *Proceedings of the 1st international conference on Embedded networked sensor systems* (pp. 266-279). ACM.
5. Wang, C., Sohraby, K., Lawrence, V., Li, B., & Hu, Y. (2006, June). Priority-based congestion control in wireless sensor networks. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, 2006.* (Vol. 1, pp. 8-pp).IEEE.
6. Hull, B., Jamieson, K., &Balakrishnan, H. (2004, November). Mitigating congestion in wireless sensor networks. In *Proceedings of the 2nd international conference on Embedded networked sensor systems* (pp. 134-147). ACM.
7. Yaghmaee, M. H., &Adjeroh, D. (2008, June). A new priority based congestion control protocol for wireless multimedia sensor networks. In *International Symposium on a World of Wireless, Mobile and Multimedia Networks, 2008. WoWMoM 2008. 2008* (pp. 1-8). IEEE.
8. Tassiulas, L. (1995). Adaptive back-pressure congestion control based on local information. *IEEE Transactions on Automatic Control, 40*(2), 236-250.
9. Sarkar, S., &Tassiulas, L. (2001). Back pressure based multicast scheduling for fair bandwidth allocation. In *INFOCOM 2001.Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies.Proceedings.IEEE* (Vol. 2, pp. 1123-1132).IEEE.
10. Tassiulas, L., &Sarkar, S. (2002). Maxmin fair scheduling in wireless networks. *In INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings.IEEE* (Vol. 2, pp. 763-772).IEEE.
11. Chen, L., Low, S. H., & Doyle, J. C. (2005, March). Joint congestion control and media access control design for ad hoc wireless networks. In *INFOCOM 2005.24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE* (Vol. 3, pp. 2212-2222).IEEE.
12. Paganini, F., Doyle, J., & Low, S. (2001). Scalable laws for stable network congestion control. In *Proceedings of the 40th IEEE Conference on Decision and Control, 2001.*(Vol. 1, pp. 185-190).IEEE.
13. Johari, R., & Tan, D. K. H. (2001). End-to-end congestion control for the Internet: Delays and stability. *IEEE/ACM Transactions on Networking, 9* (6), 818-832.
14. Shakkottai, S., Srikant, R., &Meyn, S. P. (2003). Bounds on the throughput of congestion controllers in the presence of feedback delay. *IEEE/ACM Transactions on Networking (TON), 11*(6), 972-981.
15. Bender, P., Black, P., Grob, M., Padovani, R., Sindhushyana, N., & Viterbi, A. (2000). CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users. *Communications Magazine, IEEE, 38*(7), 70-77.

Global Journal of Computer Science and Technology: G Interdisciplinary
Volume 14 Issue 5 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

Voip End-to-End Security using S/Mime and A Security Toolbox

By Md. Shahidul Islm & Md. Mahbub Rahman
Rajshahi University of Engineering & Technology, Bangladesh

Abstract- Voice Over Internet Protocol (VOIP) is a rapidly-growing Internet service for telephone communication. However, while it offers a number of cost advantages over traditional telephone service, it can pose a security threat, especially when used over public networks. In the absence of sufficient security, users of public networks are open to threats such as identity theft, man-in-the-middle attack, interception of messages/eavesdropping, DOS attacks, interruption of service and spam. S/MIME adds security to the message itself and can be used to provide end-to-end security to SIP. S/MIME can also offer confidentiality or integrity, or both, but it does not provide any anti-replay protection. However, we propose to use a unified architecture for the implementation of security protocols in the form of a security toolbox system. It will prevent an attack against anti-replay

Keywords: S/MIME, SIP, IPSec, replay attack, SDP.

GJCST-G Classification: C.2.0

Strictly as per the compliance and regulations of:

Voip End-to-End Security using S/Mime and A Security Toolbox

Md. Shahidul Islm [α] & Md. Mahbub Rahman [σ]

Abstract- Voice Over Internet Protocol (VOIP) is a rapidly-growing Internet service for telephone communication. However, while it offers a number of cost advantages over traditional telephone service, it can pose a security threat, especially when used over public networks. In the absence of sufficient security, users of public networks are open to threats such as identity theft, man-in-the-middle attack, interception of messages/eavesdropping, DOS attacks, interruption of service and spam. S/MIME adds security to the message itself and can be used to provide end-to-end security to SIP. S/MIME can also offer confidentiality or integrity, or both, but it does not provide any anti-replay protection. However, we propose to use a unified architecture for the implementation of security protocols in the form of a security toolbox system. It will prevent an attack against anti-replay.

Keywords: S/MIME, SIP, IPSec, replay attack, SDP.

I. Introduction

How can a client be sure that his message will not be intercepted by someone? This is the most important and urgent question that security professionals have to answer when dealing with VoIP systems.

Voice over Internet Protocol is a rapidly growing Internet service. Voice over IP (VoIP) has been developed in order to provide access to voice communication anywhere in the world. VoIP is simply the transmission of voice conversations over IP-based networks. Although IP was originally planned for data networking, now it is also commonly used for voice networking. While VoIP (Voice over Internet Protocol) offers a number of cost advantages over traditional telephoning, it can also pose a security threat. So watertight security is needed when using VoIP, end-to-end, especially when used on a public network. There is, however, no standard for VoIP and no general solution for VoIP security. The security of VoIP systems today is often non-existent or, in the best case, weak. As a result, hackers can easily hack.

II. Review

Several writers have taken on this or similar problems. Gupta and Shmatikov [1] investigated the security of the VoIP protocol stack, as well as SIP, SDP, ZRTP, MIKEY, SDES, and SRTP. Their investigation found a number of flaws and opportunity for replay attacks in SDES that could completely smash content protection. They showed that a man-in-the-middle attack was possible using ZRTP. They also found a weakness in the key derivation process used in MIKEY.

Niccolini et al. [2] designed an intrusion prevention system architecture for use with SIP. They evaluated the effectiveness of legitimate SIP traffic in the presence of increasing volumes of malformed SIP INVITE messages in an attack scenario.

Fessi et al. [3] proposed extensions to P2P SIP and developed a signaling protocol for P2P SIP that uses two different Kademlia-based overlay networks for storing information and forwarding traffic. Their system requires a centralised authentication server, which provides verifiable identities at the application/SIP layer.

Palmieri and Fiore [4] describe an adaptation of SIP to provide end-to-end security using digital signatures and efficient encryption mechanisms. The authors developed a prototype implementation and conducted a performance analysis of their scheme. However, one weakness of this system is that it is open to man-in-the-middle attacks.

Syed Abdul and Mueed Mohd Salman [5] developed Android driven security in SIP based VoIP systems using ZRTP on GPRS network. It communicated securely, using the GPRS data channel encrypted by using ZRTP technique. As it relies on ZRTP, it is probably vulnerable to man-in-the-middle attacks too.

Chirag Thaker, Nirali Soni and Pratik Patel [6] developed a new Performance Analysis and Security Provisions for VoIP Servers. This paper provided a performance analysis of VoIP-based servers providing services like IPPBX, IVR, Voice-Mail, MOH, Video Call and also considered the security provisions for securing VoIP servers.

III. Related Work

This paper considers a different solution, presenting a structure to assure end-to-end security by using the key management protocol S/MIME with the security toolbox system. S/MIME (Secure/Multipurpose Internet Mail Extensions) is a standard for public key encryption and signing of MIME data. S/MIME provides

Author α: Rajshahi University of Engineering & Technology (RUET) DEP: ETE, Rajshahi, Bangladesh. e-mail: sohids@zoho.com
Author σ: University of Information Technology & Sciences,(UITS) Dept: EEE,Dhaka, Bangladesh. e-mail: milueee@gmail.com

end-to-end integrity, confidentiality protection and does not require the intermediate proxies to be trusted. However, S/MIME does not provide any anti-replay protection. To protect against a replay attack, we use the security toolbox system. Toolbox system is a protocol as a single package comprised of two layers: control and a library of algorithms.

IV. Parameters' of a Solution

SIP is an application-layer protocol standardized by the Internet Engineering Task Force (IETF), and is designed to support the setup of bidirectional communication sessions for VoIP calls. The main SIP entities are endpoints (softphones or physical devices), a proxy server, a registrar, a redirect server, and a location server.

However, TLS (Transport Layer Security) can be used to introduce integrity and confidentiality to SIP between two points. Although it uses SIP signaling to secure, it has some limitations. Each proxy needs the SIP header in clear text to be able to route the message properly. All proxies in use in a connection must be trusted, as messages are decrypted and encrypted in each node. There will be no assurance that an SIP message cannot be intercepted by someone in the network.

IPSec can also be used to provide confidentiality, integrity, data origin authentication and even replay protection to SIP. It cannot be used in end-to end security. Proxy servers need to read from SIP headers and sometimes write to them. It can be used in protecting data flows between a pair of hosts (host-to-host), between a pair of security gateways (network-to-network), or between a security gateway and a host (network-to-host). IPSec assumes, however, that a pre-established trust relationship has been introduced between the communicating parties, making it most suited for SIP hosts in a VPN scenario. Further, the SIP specification does not describe how IPSec should be used; neither does it describe how key management should be operated.

S/MIME is a set of specifications for securing electronic mail and can also be used to secure other applications such as SIP. S/MIME provides security services such as authentication, non-repudiation of origin, message integrity, and message privacy. Other security services include signed receipts, security labels, secure mailing lists, and an extended method of identifying the signer's certificate(s) etc.

S/MIME provides open, interoperable protocols that allow compliant software to exchange messages that are protected with digital signatures and encryption. S/MIME requires that each sender and recipient have an X.509-format digital certificate, so public-key infrastructure (PKI) design and deployment is a major part of S/MIME deployment.

The same mechanisms can be applied for SIP. The MIME security mechanism is referred to as S/MIME and is specified in RFC 2633. S/MIME adds security to the message itself and can be used to provide end-to-end security to SIP.

Suppose two clients are trying to communicate each other. One client wants to send a message to the other client.

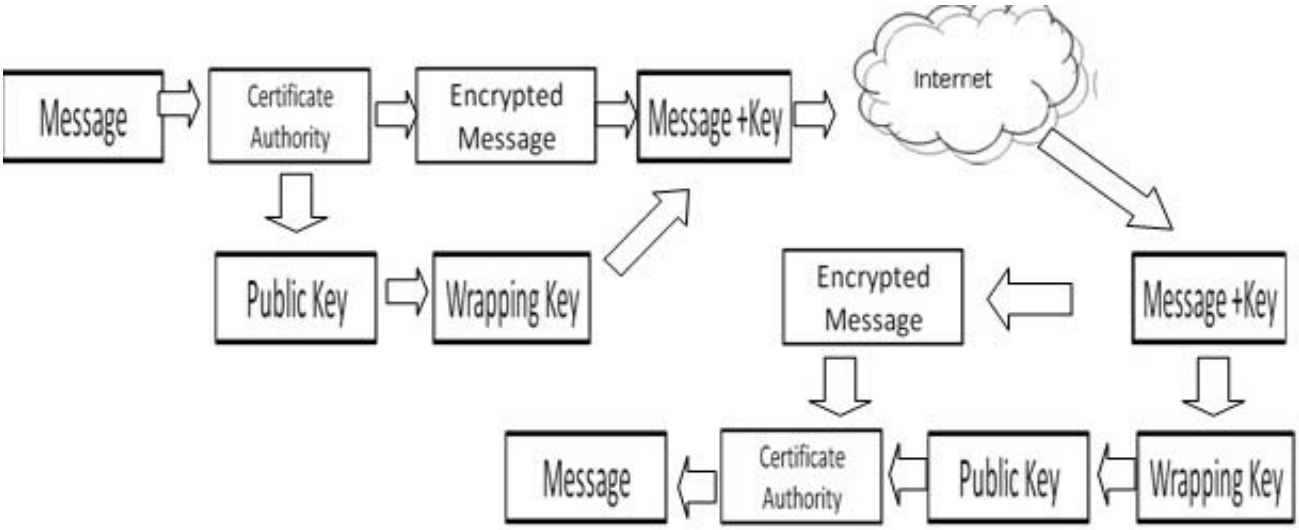


Figure 1: End-to-end security

Figure 1 : shows how to send the message in secure way.

Before S/MIME can be used to encrypt the message, one needs to obtain a key/certificate, either from one's in-house certificate authority (CA) or from a public CA.

The client uses S/MIME to sign and/or encrypt a SIP message. S/MIME combines public-key and secret-key cryptography. To encrypt the message, the sender obtains certificates from the certificate authority (CA) and generates a strong, random secret key. The message is then signed with the private key of the sender.

The encryption of the message is a bit trickier. It requires that the public key of the recipient is known to the sender. This key must be fetched in advance or be fetched from some kind of central repository. The secret key is used to encrypt the message, and then the public key of the recipient is used to encrypt the key for the recipient. When the recipient gets the message, he uses the private key to decrypt his copy of the secret key, and the secret key is used to decrypt the original message.

V. The Security Risk

S/MIME does not provide any anti-replay protection. The most serious attack is a replay attack on SDES, which causes SRTP to repeat the key stream used for media encryption, thus completely breaking transport-layer security. To protect against a replay attack, we use the security toolbox. How to use it to prevent an attack on SRTP, when used in combination with an SDES key exchange, is described below.

Suppose two users, Alice and Bob are trying to communicate with each other. Bob is the initiator in this session, and SDES is used to transport SRTP key material. To provide confidentiality for the SDES message, S/MIME is used to encrypt the payload.

S/MIME does not provide any anti-replay protection. Suppose an attacker, Charles, is trying to attack the call. Charles sends the copy of Bob's original INVITE message to Alice, containing an S/MIME-encrypted SDP attachment, with the SDES key transfer message. Since Alice does not maintain any state for SDP, she will not be able to detect the replay. Charles will effectively, for Alice, become Bob!

This is why it is proposed to use security toolbox: to prevent such a personation attack. Since anti-replay tools will be maintained all states for SDP, at all times, all messages will be filtered through anti-replay tools. Anti-replay tools will be able to detect the replay. S/MIME provides the security at the document level and IPSec performs the same function at the packet level. This configuration should become common whenever an application uses S/MIME as a document-level protection.

VI. A Security Toolbox

Ibrahim S. Abdullah and Daniel A. Menasce [9] designed a security toolbox. In the toolbox, every tool carries out a specific function such as: encryption, decryption, random number generation, integrity protection, anti-replay, and header processing.

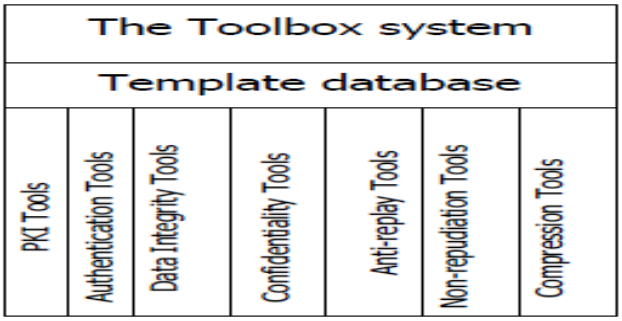


Figure 2 : Components of A Toolbox System

Figure 2 shows the major components of such a toolbox. The template is a set of specifications that define the required security services. The template database takes the necessary steps from the database for overall protection

The toolbox architecture consists of two parts: one that must be secured as part of the trusted domain of the operating system (CBT) and another that may be part of the user domain.

The secure part consists of the following components:

1. *Databases:* store information about different operations of the toolbox, such as: private and secret keys, templates, registry for the tools and template names, alert messages, authorization information, policies, and the toolbox configuration information.
2. *Interpretation engine:* interprets protocol templates.
3. *Security tools:* the set of tools that implement the security algorithms.
4. *Cache:* stores temporary keys and associated information.
5. *Inter-communication manager:* handles control messages between toolboxes running at different hosts (e.g., during handshake).

The second part of the toolbox consists of:

1. *Template developer and analyser:* analyses template creation, verification, and maintenance.
2. *Certificate repository:* contains copies of the certificates that the toolbox consults for authentication. These certificates may be placed in public storage, because they are protected by their creator's digital signature. This repository could part of a directory service application.
3. Directory services are standard applications used to provide user's authentication and authorization services.

Now let us revisit our friends Alice and Bob with a security toolbox. Recall that Bob accepts Alice's INVITE message. They communicate but then Charles sends replay messages to Alice, pretending to be Bob. Now the security toolbox takes action. First, the toolbox,

working with IPSec, has full identification of Bob: especially including his IP Packets. When Charles starts to copy Bob's messages and send them as if he were Bob, the toolbox sees that Charles' IP Packet is not the same as Bob's. Therefore, Charles is recognized as a personator and his packets are denied access to Alice. Charles' scheme fails and he goes away with nothing. Alice and Bob continue to communicate happily without any interference from hackers like Charles.

VII. Conclusion

S/MIME is being increasingly used as at security system for VoIP messages. However, S/MIME has an Achilles heel. The Achilles heel is the replay attack. This happens because S/MIME does not identify the source of the messages coming into the system. This article suggests a solution to this problem by combining the S/MIME with a security toolbox, using IPSec to monitor IP packet. The toolbox monitors the IP packet of message originators and, where a new IP address enters from the same source, denies access to the message. Such a solution guarantees complete end-to-end user security for VoIP messages at minimal cost. Thus S/MIME, with this solution, maximizes effectiveness, given the technology of the moment, in protecting the user.

References Références Referencias

1. P. Gupta and V. Shmatikov(2007), Security Analysis of Voice-over-IP Protocols. In Proceedings of the 20th IEEE Computer Security Foundations Symposium (CSFW), pages 49–63.
2. M. Petraschek, T. Hoeher, O. Jung, H. Hlavacs, and W. N. Ganstere(2008), Security and Usability Aspects of Man-in-the-Middle Attacks on ZRTP, Journal of Universal Computer Science, vol. 14, no. 5, pp. 673– 692.
3. S. Niccolini, R. G. Garroppo, S. Giordano, G. Risi, and S. Ventura(2006), SIP Intrusion Detection and Prevention: Recommendations and Prototype Implementation. In Proceedings of the 1st IEEE Workshop on VoIP Management and Security (VoIP MaSe), pages 47–52.
4. A. Fessi, N. Evans, H. Niedermayer, and R. Holz(2010), Pr2-P2PSIP: Privacy Preserving P2P Signaling for VoIP and IM, in Proceedings of the 4th Annual ACM Conference on Principles, Systems and Applications of IP Telecommunications (IPTCOMM), pp. 141–152.
5. F. Palmieri and U. Fiore(2009), Providing True End-to-End Security in Converged Voice over IP Infrastructures, Computers & Security, vol. 28, pp. 433–449.
6. Jim Murphy(2013), Toll Fraud Challenges and Prevention in a VoIP Environment (President of Phone Power).
7. Syed Abdul Mueed, Mohd Salman(2012), Android driven security in SIP based VoIP systems using ZRTP on GPRS network.(International Journal of Computer Networks and Wireless Communications (IJCNWC), ISSN: 2250-3501 Vol.2, No.2)
8. Jonathan Zar, David Endler and Dipak Ghosal(2005) VoIP Security and Privacy Threat Taxonomy (VIOPSA).
9. Ibrahim S. Abdullah and Daniel A. Menasce(2003), A unified architecture for the implementation of security protocols. In the Proc. of Computer Applications in Industry and Engineering (CAINE03), Las Vegas, Nevada USA, 11-13.

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.

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

A

Abdennadher · 32, 39
Alborzi · 9

B

Bayesian · 32

D

Dessouky · 9
Dhadesugoor · 9

E

Eltobely · 30, 40
Enqueuing · 19
Ezeife · 30, 40

G

Guennoun · 30, 40

H

Harmantzis · 29, 41

I

Ioannou · 9

K

Kademlia · 51
Kruegel · 30, 41

L

Lbekkouri · 30, 40
Lmbench · 21

M

Mahalanobis · 29, 33, 35, 36
Mazieres, · 21
Mcalerney · 30, 42
Morimura · 9

N

Niccolini · 51, 54

O

Ousterhout · 19, 21
Owerri · 11, 19

P

Palmieri · 51, 54
Perdisci · 30, 41

S

Scholkopf · 35, 41
Sebastiani · 35, 41
Shakkottai · 50
Shmatikov · 51, 54

U

Ultrasparcs · 21

Y

Yaghmaee · 49

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