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

Global Journal

of Computer Science and Technology: E

Network, Web & Security

Highlights

Data Gathering Scheme Wireless Sensor Networks

Path-Constrained Data Data-Gathering Protocol

Discovering Thoughts, Inventing Future

VOLUME 13 ISSUE 12 VERSION 1.0

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

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

Volume 13 Issue 12 (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 Inc., Headquarters Corporate Office, Cambridge Office Center, II Canal Park, Floor No. 5th, ***Cambridge (Massachusetts)***, Pin: MA 02141 United States
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Association of Research, Marsh Road, Rainham, Essex, London RM13 8EU United Kingdom.

Packaging & Continental Dispatching

Global Journals, 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: *investers@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)

Editorial Board Members (HON.)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Contents of the Volume

Global Journal of Computer Science and Technology
Network, Web & Security
Volume 13 Issue 12 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Synthesis Approach of 2D Mesh Network Inter Communication (2D-2D) using Network on Chip

By Prachi Agarwal, Dr. Anil Kumar Sharma & Adesh Kumar
University of Petroleum & Energy Studies, India

Abstract - The solution for the multiprocessor system architecture is Application specific Network on Chip (NOC) architectures which are emerging as a leading technology. Modeling and simulation of multilevel network structure and synthesis for custom NOC can beneficial in addressing several requirements such as bandwidth, inter process communication, multitasking application use, deadlock avoidance, router structures and port bandwidth. The paper emphasizes on the network on chip modeling and synthesis of 2D network and intercommunication among multilevel 2D networks. NOC synthesis environment provides transaction level network modeling and address all the requirements together in an integrated chip. In the paper consideration is done for 2D, 8 x 8 network and similar networks are considered which are identified by their specific network address. NOC chip is developed using VHDL programming language. Design is implemented in Xilinx 14.2 VHDL software, functional simulation is carried out in Modelsim 10.1 b, student edition and synthesis process is carried out on Digilent Sparten -3E FPGA.

Keywords : network on chip (NOC), very high speed integrated circuit hardware description language (VHDL), field programmable gate array (FPGA), application specific integrated circuit (ASIC).

GJCST-E Classification : C.2.1

Strictly as per the compliance and regulations of:

Synthesis Approach of 2D Mesh Network Inter Communication (2D-2D) using Network on Chip

Prachi Agarwal [α], Dr. Anil Kumar Sharma [σ] & Adesh Kumar [ρ]

Abstract - The solution for the multiprocessor system architecture is Application specific Network on Chip (NOC) architectures which are emerging as a leading technology. Modeling and simulation of multilevel network structure and synthesis for custom NOC can beneficial in addressing several requirements such as bandwidth, inter process communication, multitasking application use, deadlock avoidance, router structures and port bandwidth. The paper emphasizes on the network on chip modeling and synthesis of 2D network and intercommunication among multilevel 2D networks. NOC synthesis environment provides transaction level network modeling and address all the requirements together in an integrated chip. In the paper consideration is done for 2D, 8 x 8 network and similar networks are considered which are identified by their specific network address. NOC chip is developed using VHDL programming language. Design is implemented in Xilinx 14.2 VHDL software, functional simulation is carried out in Modelsim 10.1 b, student edition and synthesis process is carried out on Digilent Sparten -3E FPGA.

Keywords : network on chip (NOC), very high speed integrated circuit hardware description language (VHDL), field programmable gate array (FPGA), application specific integrated circuit (ASIC).

I. Introduction

Network on Chip (NoC) [1] [2] [3] is the latest approach to overcome the limitation of bus based communication network. NoC is a set of routers employed in a network, in which different nodes are inter connected with their cores can communicate with each others. In a network data comes in packets and sent to the destination with IP via routers and links [4] . When a packet reaches its destination address, it means it is switched [5] to the IP attached to the router. On-chip communications among different networks is possible using interconnection network topology [5] [6], switching, routing, queuing .flow control [11] and scheduling. Research can be done for n- dimensional topological structures network on chip design. The idea of NoC is derived from distributed computing and large scale computer networks. There are different routing techniques used in NOC design considerations to meet high throughput and cover time to market. Due to big constraints on hardware and memory resources utilization, the routing methods for NoC should be very simple.

According to the need of processors, NoC technology gives chip designer's flexibility in choosing the network topology, according to University of Bologna professor Luca Benini, founder of and scientific advisor for iNoCs, a start-up provider of on-chip interconnection technology. High degree of parallelism and pipelining increases [8] [12] the performance of the system because all the links in the network works simultaneously on different data packets [13]. As the complexity of the system is increasing, NOC is the solution to enhance system performance in comparison to the previous technological architectures such as dedicated, wires, point to point, bridges and shared buses. The scalability of system and throughput [17] [20] will increase because algorithms are designed in such a way that it offers higher degree of parallelism. For example, a mesh NoC topology [18] can function with parallelism and thus is well-suited for multiprocessor SoCs, whose cores must run in parallel. Prototype NoCs by the Electronics and Information Technology Laboratory of the French Atomic Energy Commission's Faust, the Swedish Royal Institute of Technology's Nostrum, and the Technion-Israel Institute of Technology's QNoC work with a mesh topology.

a) Tools Utilized

Design and implementation of mesh network is carried out using Project Navigator ISE 14.2, Xilinx company. It is a tool used to design the IC and to view their RTL (Register Transfer Logic) schematic. Model Sim EE 10.1b student's edition is a tool of Mentor Graphics Company used for simulation and debugging the functionality. The chip implementation is done using VHDL programming language.

The paper is organized as follows: Section I presents the introduction and the tools utilized. Section II describes intercommunication among 2D (8 x 8) mesh networks. Section III describes the FPGA synthesis environment. Section IV describes the Result and Performance Evaluation. Section V presents the Device utilization and timing summary. Conclusion is presented in Section VI.

Author α : M.Tech Scholar, Department of Electronics, Institute of Engineering & Technology, Alwar, Rajasthan India.
E-mail : prachiagarwal_05@yahoo.co.in
Author σ : Professor, Department of Electronics, Institute of Engineering & Technology, Alwar, Rajasthan India. E-mail : aks_826@yahoo.co.in
Author ρ : Assistant Professor, Department of Electronics Engineering, University of Petroleum & Energy Studies, Dehradun India.
E-mail : adeshmanav@gmail.com

II. Intercommunication among 2D Mesh Networks

Intercommunication among 2D NOC networks can be understood using arbitration logic selection of networks [12]. First understanding, how a 2D networks behaves, then focusing on the intercommunication among 2D networks. 2D NOC [12] follows the cross link which allows addressing any node at any time [11]. A 2D mesh network is connecting multiple inputs to multiple outputs in a matrix form. The 2D NOC architecture is a m × n mesh of switches [10] and resources are placed on the slots formed by the switches. For an m x n architecture there are m nodes on X axis and n nodes on Y axis respectively. Considering an 8 x 8 structure in which 64 nodes can perform intra communication. Node identification is based on the row address and column address [1]. For example, if row address = 000 and column address = 101, node 6 (N_6) is identified. Similarly there is the possibility of identifying any node. Table 1 list the possible node address generation scheme for 2D 8 x 8 structure.

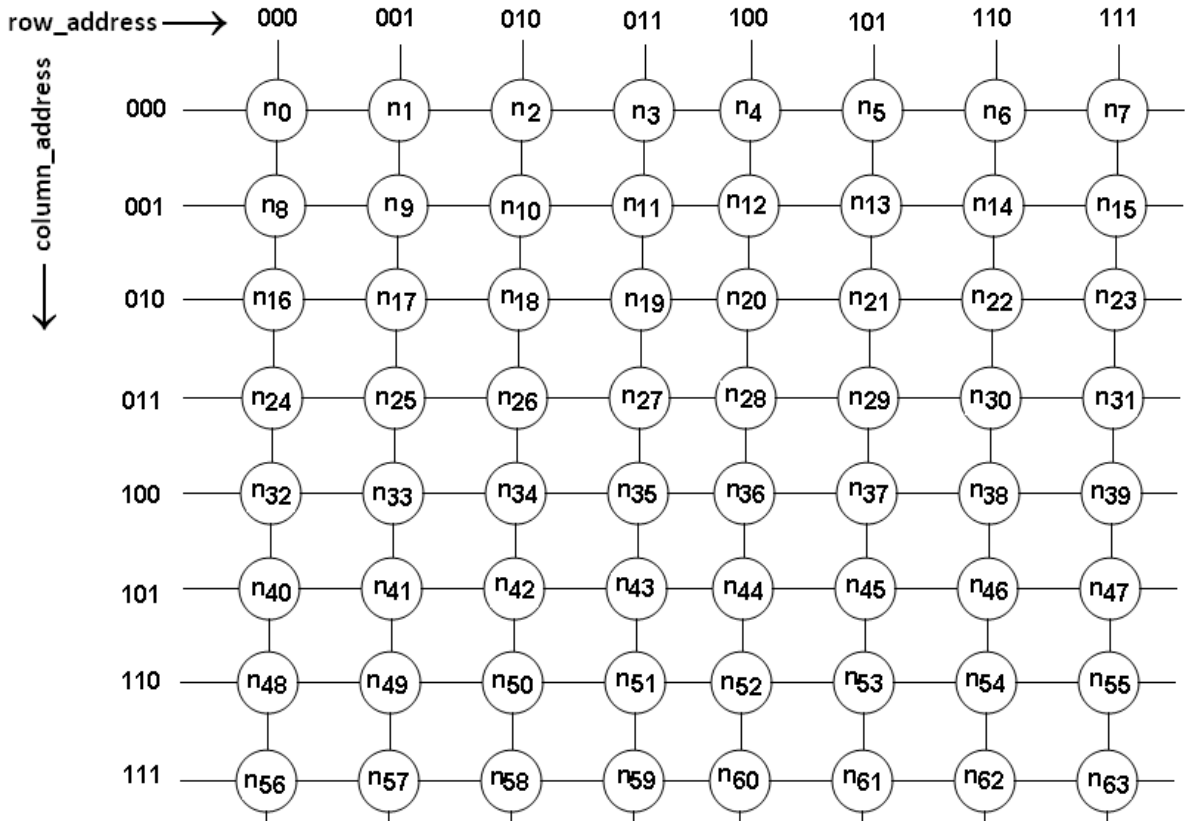

Figure 1 : Two dimensional (2D) cross point topological (8 × 8) structure [12]

Table 1 : Node address generation scheme in 2D structure

Row Address	Column Address							
	000	001	010	011	100	101	110	111
000	n_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
001	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}
010	N_{17}	N_{18}	N_{19}	N_{20}	N_{21}	N_{22}	N_{23}	N_{24}
011	N_{25}	N_{26}	N_{27}	N_{28}	N_{29}	N_{30}	N_{31}	N_{32}
100	N_{33}	N_{34}	N_{35}	N_{36}	N_{37}	N_{38}	N_{39}	N_{40}
101	N_{41}	N_{42}	N_{43}	N_{44}	N_{45}	N_{46}	N_{47}	N_{48}
110	N_{49}	N_{50}	N_{51}	N_{52}	N_{53}	N_{54}	N_{55}	N_{56}
111	N_{57}	N_{58}	N_{59}	N_{60}	N_{61}	N_{62}	N_{63}	N_{64}

Considering a architecture in which four 2D (8 x 8) can configured in such a way that they can communicate each other. Each network is identified using its network_address. If network_address is 00 Network 1, network_address is 01 Network 2, network_address is 10 Network 3, and network_address is 11 Network 4 is identified. It is also illustrated using figure 3 and table 2.

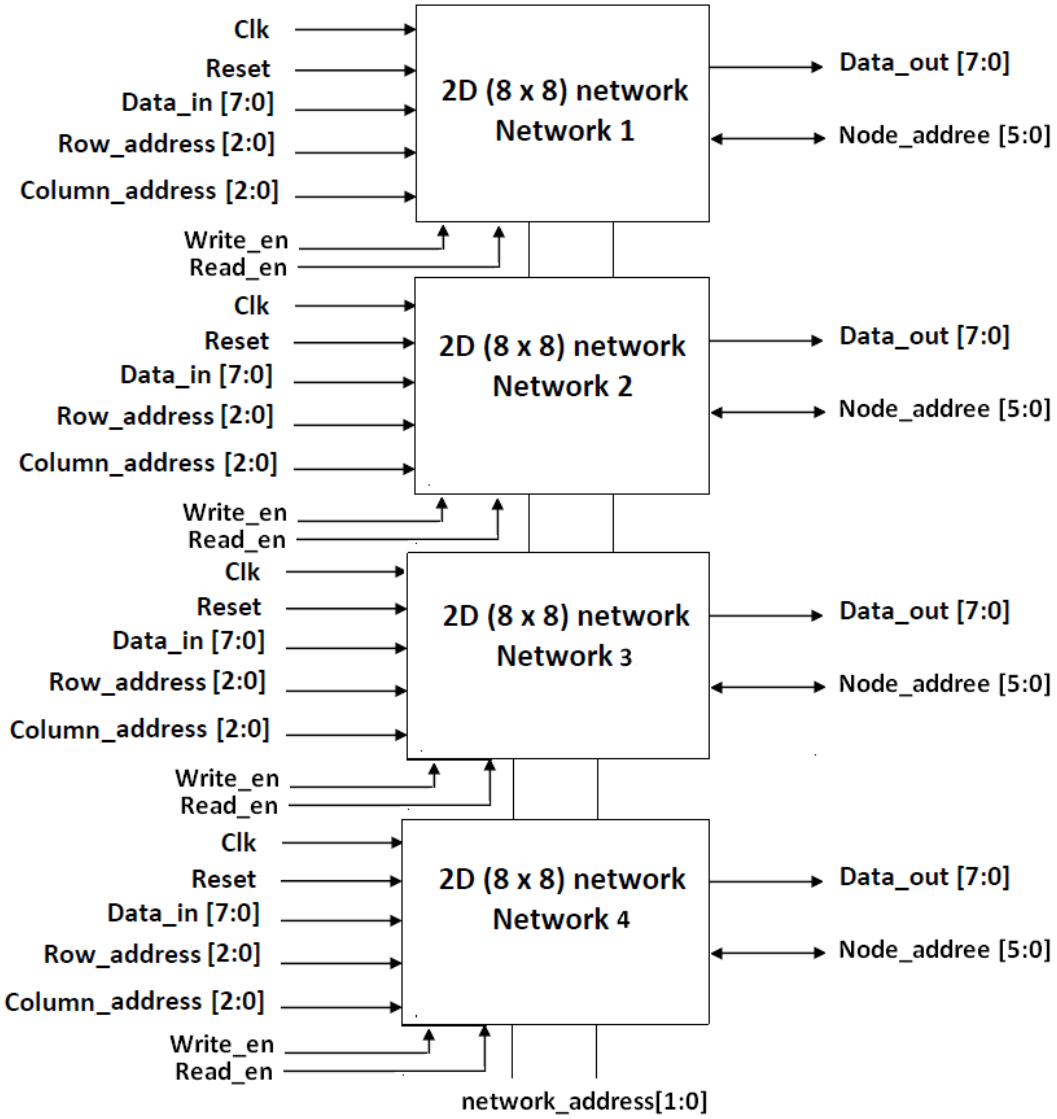

Figure 2 : Intercommunication Mesh Network

Table 2 : Network Selection

Network_Address	Selection Logic
00	Network 1 is selected
01	Network 2 is selected
10	Network 3 is selected
11	Network 4 is selected

III. Result & Performance Evaluation

Figure 3 shows the simulated result for the 8 x 8 intercommunication architecture, which shows 8 bit data transfer for network 1 to network 4. The functional simulation depends on the steps and Modelsim output is extracted after completion of these steps.

a) Simulation Process Sequence

Step 1: reset = 1, clk is used for synchronization and then run.

Step 2: reset = 0, same *clk* is used for synchronization and provide rising edge.

Step 3: Select the address of destination node *Node_address [5:0]* of 6 bits for 8 x 8 structure.

Step 4: Force the value of *row_address* and *column_address* of destination node. For 8 x 8 NOC *row_address[2:0]* and *column_address[2:0]* are of 3 bits.

Step 5: Force the value of *network address [1:0]* of destination network. *network_address [1:0] = "00"* for network 1, *network_address [1:0] = "01"* for network 2, *network_address [1:0] = "10"* for network 2 and *network_address [1:0] = "11"* for network 3.

Step 6: Give the eight bit value of data_in. Force write_en =1 and read_en =0 and then run.

Step 7: Write_en =0 and read_en =1 and run. Desired output on destination is achieved.

When write_en =1 and read_en =0, the data is written in temp variable from the source node, when write_en =0 and read_en =1, the data is read from the temp variable to destination node. Clk is applied at the

positive edge clock pulse and reset is kept at 1 for the initial state. The Register transfer level (RTL) view of chip is shown yin the figure 4 and the details of each pin is listed in table 3.

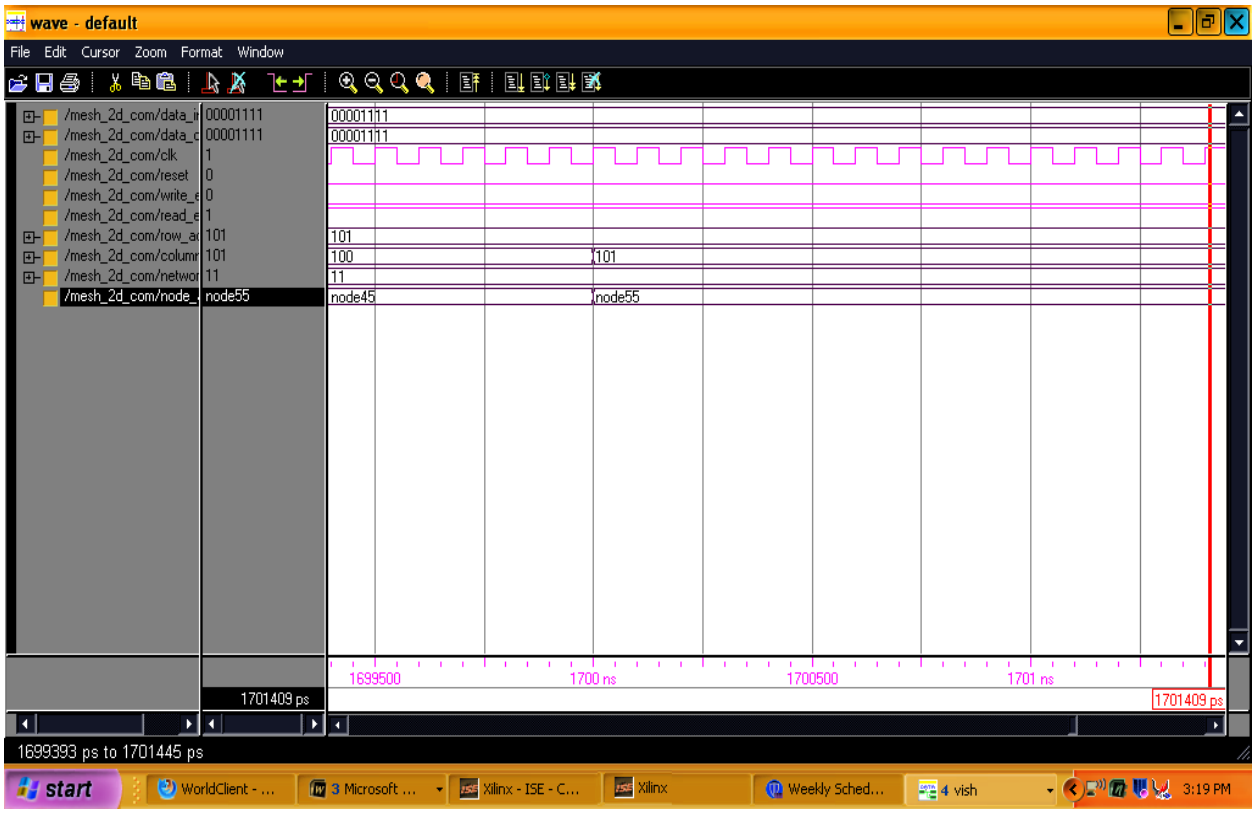

Figure 3 : Modelsim Output of Mesh Intercommunication Network (2D-2D) (8 x 8)

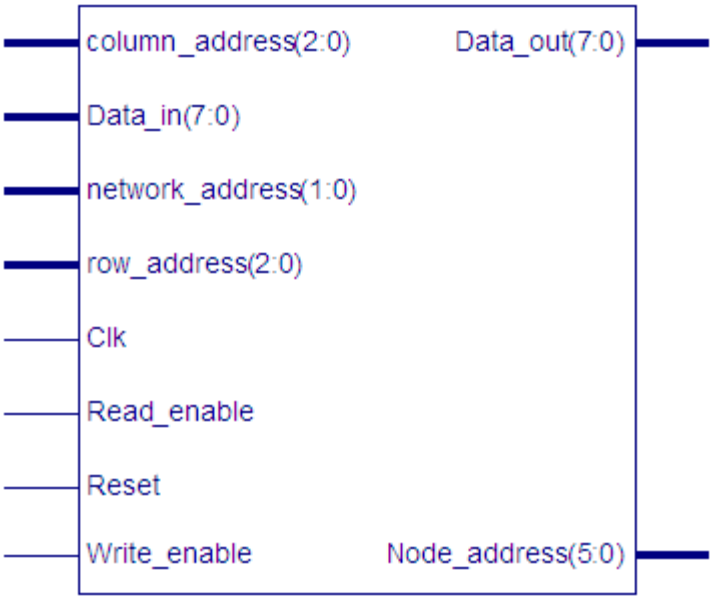

Figure 4 : RTL view of 2D-2D chip

Table 3 : Design pins and their functional description for (8 x 8) NOC

Pins	Functional Description
reset	used for synchronization of the components by using clk
clk	Provide rising edge of clock pulse
node_address [5:0]	Address of the source and destination node of 6 bits
row_address [2:0]	represents address of the nodes in x direction (3 bits)
column_address [2:0]	represents address of the nodes in y direction (3 bits)
read_en	control signal to read data (1 bit)
write_en	control signal to write data (1 bit)
network_address [1:0]	Selection logic for the network (2 bits)
data_in[7:0]	represents input data in the network (8 bits)
data_out[7:0]	represents 8 bit output data of the destination node (8 bits)

IV. FPGA Synthesis

The Spartan-3E starter kit [21] [22] provides easy way to test the various programs in the FPGA itself, by dumping the 'bit' file of the designed program in Xilinx software into the FPGA and then observing the output .The Spartan 3E FPGA board [8] comes built in with many peripherals that help in the proper working of the board and also in interfacing the various signals to the board itself. Some of the peripherals included in the Spartan 3E FPGA board include: 2-line, 16-character LCD screen used for display the output, PS/2 mouse or keyboard port can be connected to the FPGA, VGA display port [21] used to display various encoded data on screen. The Sparten 3E kit is shown in the figure 5. Switches are the input for *clk, reset*, *row_addressn[2:0]*, *column_address [2:0]*, *network_address [1:0]*. Eight bits data *data_in[7:0]* is also given using switches. These switches are locked into FPGA using user constraint file (UCF). Figure 6 shows the flow of inputs given to FPGA device. Four slide switches and four push-button switches are used to give the inputs to the FPGA board. They can also act as the reset switches for the various program Kit also has four-output, SPI-based on board Digital-to-Analog Converter (DAC) on board which is to be interfaced to give the analog output to the digital data values. Two-input, SPI-based [7] [23] Analog-to-Digital Converter (ADC) with programmable gain preamplifier converts the real world analog signals into digital values.' Image processing inputs are given by the switches of kit and functionally tested on the corresponding LED's output.T he output data is flashed on LEDs. These LEDs are also locked in UCF file [16]. The bit file of the program is burn out in The EPROM of FPGA and corresponding result is shown by blinking LEDs. The output can be shown on Digital storage oscilloscope (DSO). As shown in figure 7. The input data is 10101010, when reset switch = 0, no output is display on DSO. When reset switch =1, output data is 10101010.

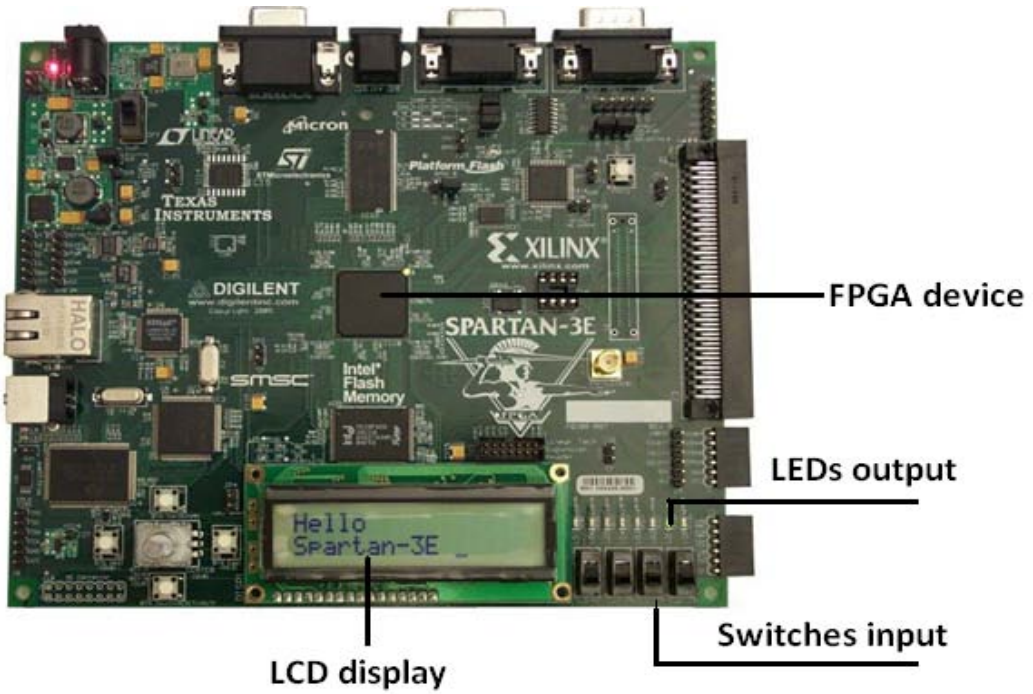

Figure 6 : Sparten-3 FPGA view [21]

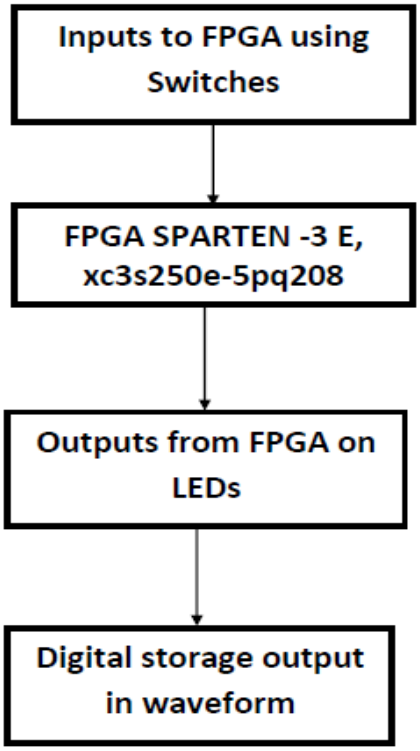

Figure 7 : FPGA syntheis flow on Sparten -3E

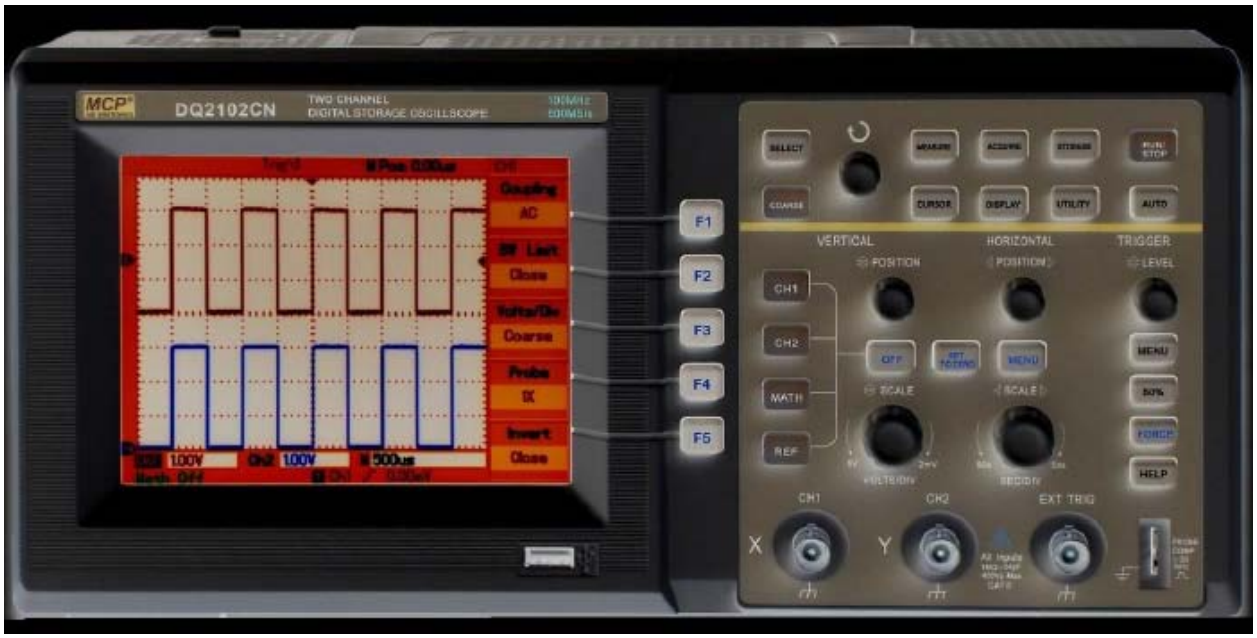

Figure 8 : Synthesis Output on DSO

V. Device Utilization and Timing Summary for 2D NOC

Device utilization report gives the percentage utilization of device hardware for the chip implementation. Timing report generates minimum and maximum time [1] [12] to reach the output. Synthesis report [21] extracted from the Xilinx shows the complete details of device utilization and timing summary.

Selected Device: xc3s250e-5pq208, this device is targeted for FPGA. Device utilization summary for 8 x 8 2D Mesh NOC is shown in table 5 and 6 respectively.

Table 5 : Device utilization in (2D-2D) Mesh structure

Logic Utilization	Used	Available	Utilization
Number of Slices	335	2448	13 %
Number of Slice Flip Flops	80	4896	1 %
Number of 4 input LUTs	653	4896	13 %
Number of bonded IOBs	31	158	19 %
Number of GCLKs	2	24	8 %

a) Timing Summary for 8 x 8 NOC

Timing details provides the information of delay, minimum period, minimum input arrival time before clock and maximum output required time after clock [1].
Speed Grade: - 5
Minimum Period : 4.937 ns (Maximum Frequency: 202.536 MHz).
Mininput arrival time before clock : 9.977 ns.
Max output required time after clock : 4.368 6ns.
Total memory usage is 127274 kilobytes.

VI. Conclusion

The hardware chip for 2D-2D intercommunication network is modeled and simulated in Xilinx 14.2 successfully. The network size is considered 8 x 8 in which 8 nodes can communication among each other in duplex mode. The data communication from one network to another network is tested for different test cases. The network structure is synthesized in Xilinx supporting Digilent Sparten-3E FPGA kit. Device utilization summary shows Slices utilization 13 %, Number of Slice Flip Flops 1 %, Number of 4 input LUTs 13 %, Number of bonded IOBs 19 % and Number of GCLKs 2 %. Memory utilization is 127274 kB and maximum frequency 202.536 MHz. These parameters are optimized parameters. The proposed architecture is applicable for inter and intra communication among networks. In future we can enhance the data size, and number of nodes supporting to the networks.

References Références Referencias

1. Prachi Agarwal, Anil Kumar Sharma, Adesh Kumar "Modeling and Simulation of 2D Mesh Topological Network on Chip (NOC)" International Journal of Computer Applications (0975 – 8887) Volume 72– No.21, June 2013 page (25-30).
2. B. Grot, J. Hestness, S. W. Keckler and O. Mutlu. Express Cube Topologies for On-Chip Interconnects. In *15th International Symposium on Computer Architecture (HPCA)*, 2009.
3. David Atienzaa, Federico Angiolini, Srinivasan Murali, Antonio Pullinid, Luca Benini, Giovanni De Micheli, Network-on-Chip design and synthesis outlook, INTEGRATION, the VLSI journal Elsevier 41 (2008) 340–359.
4. Davide Bertozzi and Luca Benini, A Network-on-Chip Architecture for Gigascale Systems-on-Chip, IEEE Circuits and systems magazine, second quarter 2004, Xpipes, page (2-6).
5. F. Karim A. Nguyen, and S. Dey, "An interconnect architecture for networking systems on chips", IEEE Journal on Micro High Performance Interconnect, vol. 22, issue 5, pp. 36-45, Sept 2002.
6. Jason Cong, Yuhui Huang, and Bo Yuan Computer Science Department University of California, Los Angeles Los Angeles, USA, 978-1-4577-1400-9/11/$26.00 ©2011 IEEE, A Tree-Based Topology Synthesis for On-Chip Network, pp 2-6.
7. J. D. Owens, W. J. Dally etal., "Research challenges for on-chip inter connection networks," IEEE MICRO, vol. 27, no. 5, pp. 96–108,Oct. 2007.
8. H. G. Lee, N. Chang, U. Y. Ogras, and R. Marculescu, "On-chip communication architecture exploration: A quantitative evaluation of point-to-point, bus, and network-on-chip approaches," ACM Trans. Des. Autom. Electron. Syst., vol. 12, no. 3, pp. 1–20, Aug. 2007.
9. Mohammad Ayoub Khan, Abdul Quaiyum Ansari, A Quadrant-XYZ Routing Algorithm for 3-D Asymmetric Torus Network-on-Chip, The Research Bulletin of Jordan ACM, ISSN: 2078-7952, Volume II (II) pp(18-26).
10. M. Coppola, S. Curaba, M. Grammatikakis, R. Locatelli, G. Maruccia, F. Papariello, L. Pieralisi, White paper on OCCN: A Network-On-Chip Modeling and Simulation Framework, ISD Integrated system developments, page 8.
11. Naveen Chaudhary, Bursty Communication Performance Analysis of Network-on-Chip with Diverse Traffic Permutations International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-6, January 2012, (page 1).
12. Adesh Kumar, Sonal Singhal, Piyush Kuchhal "Network on Chip for 3D Mesh Structure with Enhanced Security Algorithm in HDL Environment" International Journal of Computer Applications (IJCA), USA, (ISBN: 973-93-80871-97-9) Volume 59– No.17 (page 6-13).
13. P. Pratim Pande, C. Grecu, M. Jones, A. Ivanov, and R. Saleh, "Performance evaluation and design trade-offs for network-on-chip interconnect architectures", IEEE Transactions on Computers, vol. 54, no. 8, pp. 1025-1040, 2005.
14. Rikard Thid Thesis on "A Network on Chip Simulator", Sweden Master of Science Thesis in Electronic System Design, Royal Institute of Technology Aug 2002, Page (9-27).
15. S. Borkar, "design challenges of technology scaling." IEEE Micro, no.4, p. 2329, July-August 1999.
16. Semiconductor Complex Limited, Internet PDF: Data sheets of XC 95 series CPLD [Online].

17. Vitorde Paulo and Cristinel Ababei, 3D Network-on Chip Architectures Using Homogeneous Meshes and Heterogeneous Floorplans, Hindawi Publishing Corporation International Journal of Reconfigurable Computing Volume 2010, Article ID603059, 12 pages.
18. W. Wolf, The future of multiprocessor systems-on-chips, in: Proceedings of the 41st Design Automation Conference (DAC'04), June 2004, pp. 681–685.
19. Woo-seo Ki1, Hyeong-Ok Lee, Jae-Cheol Oh ,"The New Torus Network Design Based On 3-Dimensional Hypercube", ICACT, pp. 615-620, Feb.15-18 2009.
20. Xin Wang and Jari Nurmi, Comparison of a Ring On-Chip Network and a Code-Division Multiple-Access On-Chip Network, Hindawi Publishing Corporation VLSI Design Volume2007, ArticleID 18372, 14 page.
21. http://www.xilinx.com/support/documentation/data_sheets/ds312.pdf
22. http://www.xilinx.com/support/documentation/data_sheets/ds099.pdf

Global Journal of Computer Science and Technology
Network, Web & Security
Volume 13 Issue 12 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Evaluating Smartphone Application Security: A Case Study on Android

By Muneer Ahmad Dar & Javed Parvez
University of Kashmir, India

Abstract - Currently, smart phones are becoming indispensable for meeting the social expectation ofalways staying connected and the need for an increase inproductivity are the reasons for the increase in smartphone usage. One of the leaders of the smart phone evolution is Google's Android operating system. It ishighly likely that Android is going to be installed in manymillions of cell phones during the near future. With thepopularity of Android smart phones everyone finds it convenient to make transactions through these smartphones because of the openness of Android applications. The malware attacks are also significant. Androidsecurity is complex and we evaluate an applicationdevelopment environment which is susceptible tomalware attacks. This paper evaluates Android securitywith the purpose of identifying a secure applicationdevelopment environment for performing securetransactions on Android-based smart phones.

Keywords : *smartphone, android, malware, spam, vulnerabilities, attacks, mobility, API.*

GJCST-E Classification : *D.4.6*

Strictly as per the compliance and regulations of:

Evaluating Smartphone Application Security: A Case Study on Android

Muneer Ahmad Dar [α] & Javed Parvez [σ]

Abstract - Currently, smart phones are becoming indispensable for meeting the social expectation of always staying connected and the need for an increase in productivity are the reasons for the increase in smart phone usage. One of the leaders of the smart phone evolution is Google's Android operating system. It is highly likely that Android is going to be installed in many millions of cell phones during the near future. With the popularity of Android smart phones everyone finds it convenient to make transactions through these smart phones because of the openness of Android applications. The malware attacks are also significant. Android security is complex and we evaluate an application development environment which is susceptible to malware attacks. This paper evaluates Android security with the purpose of identifying a secure application development environment for performing secure transactions on Android-based smart phones.

Keywords : smartphone, android, malware, spam, vulnerabilities, attacks, mobility, API.

I. Introduction

Smartphones have become indispensable part of our daily lives in recent years, since they are involved in keeping in touch with friends and family, doing business, accessing the internet and other activities. Andy Rubin, Google's director of mobile platforms, has commented: "There should be nothing that users can access on their desktop that they can't 0access on their cell phone" [1]. Growth in smartphone sales is depicted in the figure below.

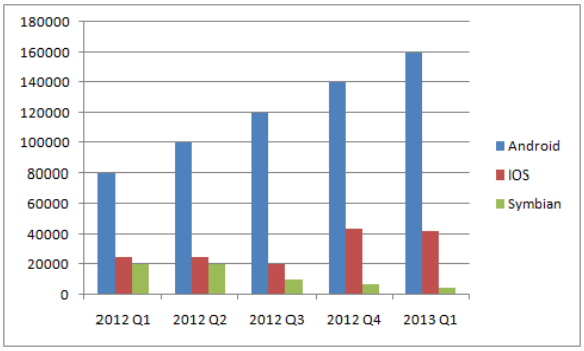

Figure 1 : Worldwide Smartphone Sales

Author α : Scientist-B at National Institute of Electronics and Information Technology Srinagar, India, received Masters degree in Computer Applications from the University of Kashmir.
E-mail : muneerdar07@gmail.com
Author α : Senior Assistant Professor at University of Kashmir, received a BE in Electrical & Electronics Engineering from BITS-Pilani, India, MS in Computer Science from University of Oklahoma (USA) and PhD in Computer Science from the University of Kashmir, India.
E-mail : javedparvez1225@gmail.com

It indicates that smart phone sales are continuously on rise and more and more people are becoming dependent on these devices. As these Smartphones are going to outnumber the world's total population in 2014, securing these devices has assumed paramount importance. Owners use their smart phones to perform tasks ranging from everyday communication with friends and family to the management of banking accounts and accessing sensitive Work related data. These factors, combined with limitations in administrative device control through owners and security critical applications like the banking transactions, make Android-based Smart phones a very attractive target for hackers, attackers and malware authors with almost any kind of motivation.

In this paper we analyze the security architecture of Smartphones in section II and identify the security loopholes of the current framework in section III. We then have a detailed description of the existing work done by the researchers in section IV. Finally we draw our conclusions in section V.

II. Security Framework of Smartphone

A Smartphone is an intricate combination of a mobile phone and a computing platform, with high-speed connectivity and powerful computing ability. Therefore, the Smartphone has necessary components of the computing platform: an operating system, applications and hardware. Furthermore, as a personal communication device, the Smartphone also often has multiple communication capabilities and ability to store large amounts of sensitive user data [2].

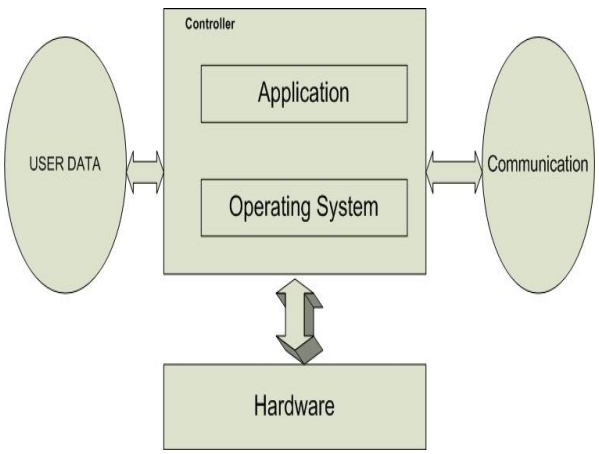

Figure 2 : Common Smartphone Architecture [2]

Android is a Linux-based platform programmed with Java and enhanced with its own security mechanisms tuned for a mobile environment. Android combines OS features like efficient shared memory, preemptive multi-tasking, Unix user identifiers (UIDs) and file permissions with the type-safe Java language and its familiar class library. The resulting security model is much more like a multi-user server than the sandbox found on the J2ME or Blackberry platforms. Unlike a desktop computer environment where a user's applications all run under the same UID, Android applications are individually siloed from each other.

Android applications run in separate processes under distinct UIDs each with distinct permissions. Programs can typically neither read nor write each other's data or code, and sharing data between applications must be done explicitly. The Android GUI environment has some novel security features that help support this isolation [10].

III. Limitations of Current Security Model

Android is based on open source framework and it comes with a pre-built suite of applications like dialer, address book, browser, etc. Developers can code their own applications and publish to the Android market after a self-signing phase that does not require any certifying authority i.e., developer can use self-created certificates to sign their applications. This helps in providing a wide range of applications and services to the device-holders. Since there is no support for root Certification Authorities in Android (Android is based on open source framework, while introducing root certification mechanisms contradicts the openness of Android) it is very difficult to scrutinize and/or block applications coming from unreliable sources. There is a higher possibility that an average Android user will easily be tricked by unsafe applications and will avoid the warning messages at time of installation.

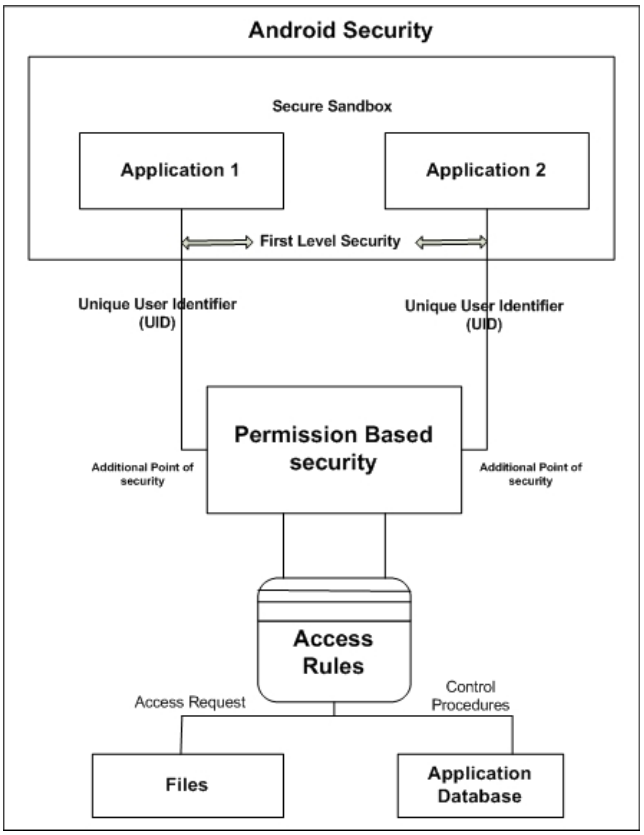

Figure 3 : Security Architecture of Android

The permission model is the core mechanism for securing access to various resources in Android. Although the permissions are categorized to different protection levels such as Normal, Dangerous, Signature and Signature-Or-System but the assignment of these protection levels is left to the developer's will and own understanding. This leads to a number of vulnerabilities in the permission model. When an application is installed on Android, the Android framework prompts the user with a list of required permissions, the user may grant all of the permissions in order to install the application or deny the permissions to decline the installation. Practically, there are a number of issues in such a model: 1) The user must grant all of the required permissions in order to install the application, 2) once the permissions are granted; there is no mechanism for further restricting an application to use the granted permissions, 3) there is no way of restricting access to the resources based on dynamic constraints as the permission model is based on install-time check only, 4) granted permissions can only be revoked by uninstalling the application. There are four security loopholes that endanger the information stored in a Smartphone [18]:

- The capability to sniff the data transmitted by any network where the device is connected.
- The lack of strong security control of user´s private information that permits malware to access the information stored in the device.

- Saved information is not stored encrypted within media.
- The lack of configurable firewalls integrated into the operating systems.

These are the most common security problems in the smart phones. Powerful hardware and advanced operating system with flexible APIs [23], [24] not only increase capability and functionality of smartphones but also present rising security threats to smartphones. Other features of smartphones also exacerbate threats to smartphones: higher bandwidth not only accelerates the Internet access, but also speeds up virus transmission; multiple peripheral interfaces not only increase Smartphone connectivity, but also provide much more avenues for virus injection. Compared with personal computers, smartphones also have always-on & always-connected mobility and are therefore hard to reinstall. Therefore, disabling or making smartphones nonfunctional as a result of attacks will have much greater impact and prove more costly than in the case of personal computers. A threat means potential violation of Smartphone security, which can be clustered into two categories according to its cause: vulnerabilities [5] and attacks [6]. Vulnerabilities mean weakness of smartphones and inabilities to withstand hostile environment effects. Attacks are any attempts to intentional destroy unauthorized use, maliciously modify or illegally obtain Smartphone assets. Generally, attackers bypass or exploit deficiencies of Smartphone security mechanisms to initiate negative activities. That is, **vulnerabilities** are the internal attributes of smartphones, while **attacks** are the outside offensive activities to smartphones. As a matter of fact, most attacks exploit vulnerabilities of smartphones.

a) Vulnerabilities

As comparatively complicated devices, smartphones inevitably have numerous vulnerabilities, which can lead to insecurity or be exploited by malicious persons to initiate attacks. Smartphone vulnerabilities typically include system defects, insufficient management of APIs, deficiency of user awareness and unsecure wireless channels, shown in Fig.4.

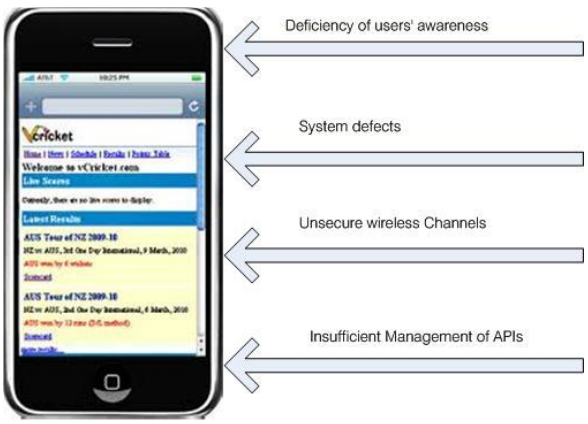

Figure 4 : Vulnerabilities of Smartphones

i. *Deficiency of User Awareness*

Some applications are installed without user confirmation or with limited information. Based on these situations, attackers can give deceptive information for applications infected by malicious codes. Users may install them without accurate or sufficient information. In addition, some sensitive operations, such as sending and receiving messages, deleting important files, activating wireless interfaces, can also be executed secretly. Consequently, Smartphone users cannot know about occurrences of malicious security-sensitive operations until the negative effects start appearing.

ii. *System Defects*

It is nearly impossible to detect and rectify all defects in Smartphone hardware and software. Some immediately non-conforming defects can be observed soon, but most other defects cannot be found for a certain time. Even if they are discovered, these defects, especially hardware defects are hard to be remedied. As a result, system defects can cause Smartphone abnormalities and malfunctions. Furthermore, malicious persons generally take advantage of existing system defects to initiate attacks and compromise Smartphone systems.

iii. *Unsecure Wireless Channels*

In wireless environments including cellular networks, user data and control signals transmitted between smartphones and network devices can be easily captured. If these data and signals are compromised, the transmitted information will be exposed.

iv. *Insufficient API Management*

The most distinct characteristic of smartphones is flexible APIs, which are used for application development [22] and installation. However, insufficient API management is also the main reason for malicious codes. Generally, Smartphone APIs are clustered into open APIs for third-party application developments and controlled APIs for remote maintenance. Controlled APIs have higher privileges, which can be used for remote system update, file erasure and information retrieval. If malicious persons obtain controlled APIs, they can initiate negative activities such as backdoor attacks. Even some open APIs may have inappropriate privileges so that they might be utilized to acquire certain privileges and initiate attacks.

b) Attacks

It is well known that smartphones have much valuable user data, especially **financial data** and **identification information**. Driven by economic benefits, many hackers have focused on smartphones and initiated multifarious attacks, shown in Fig.3:

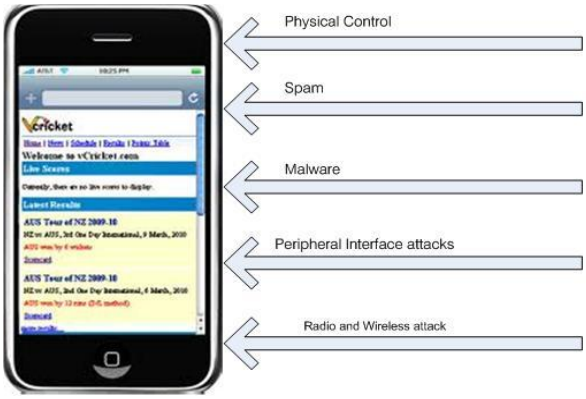

Figure 5 : Attacks to Smartphones

i. *Physical Control*

Due to portability and mobility, smartphones are likely to be lost or stolen. Then, sensitive information stored in smartphones including address book, communication records, usernames and passwords, etc. can be accessed directly. Incorrect disposals of old and damaged devices will cause similar problems.

ii. *Spam*

Spam is generally sent in SMS, MMS and email. VoIP and Instant Messaging (IM) have also become common ways for spamming. Spam may generally cause virus infection, economic loss, or worse influences.

iii. *Malware*

Malware [4], [5] is becoming the main threat to smartphones. Flexible APIs not only enrich application development, but also facilitate malware development. Meanwhile, powerful connectivity also aggravates malware spreading. In addition, smartphones can be infected by malicious codes during synchronization with personal computers or virus-infected storage media. Furthermore, malware can also be spread in a variety of ways, including Internet downloading, messaging services and Bluetooth communications. In reality, users are not always aware of downloaded applications' functions. *Even if applications have acquired explicit user consent, users may be unaware that the applications are executing malicious code*.

iv. *Backdoor*

Backdoor attacks mainly result from system bugs and disclosure of controlled APIs. Some operating systems have security loopholes such as insufficient authentication and inappropriate privileges. Based on these vulnerabilities, attackers can bypass security policies to access smartphones. In addition, if attackers have controlled APIs, they can also access smartphones like legitimate entities.

v. *Peripheral Interfaces Attacks*

Smartphones usually have many peripheral interfaces, such as Wi-Fi, Bluetooth, USB, etc. While peripheral interfaces can increase smartphones communication capabilities, unfortunately, they also become a popular steppingstone for outside attacks.

vi. *Radio and Wireless Attacks*

Due to the openness of wireless communications, attackers can easily initiate wireless attacks, which can be clustered into two categories: **active attacks** (spoofing, corruption, blocking, modifying, etc.) and **passive attacks** (sniffing, eavesdropping, etc.). Generally, passive attacks are used as a prelude to active attacks, by acquiring necessary information such as addresses and to identify vulnerabilities of potential targets.

IV. Related Research Work

a) Kirin

Enck, et al. [22] have proposed a framework known as Kirin – install-time certification mechanism – that allows the mobile device to enforce a list of pre-defined security requirements prior to installation process of an application. During installation of an application the Android framework informs the user regarding the resources that can be accessed by the application but it cannot reflect the possibility of using different combinations of permission in a malicious manner. The Kirin framework is contacted when installation process for an application package is initiated. Kirin utilizes the application's manifest file where all the required permissions are listed and uses the action string along with the permissions to construct a set of Prolog facts Although Kirin is one of the first security policy extension for Android platform, it suffers from the common limitation of Smartphone security systems i.e. the policy expressibility is not sufficient enough to express certain policies. Furthermore, the policies used by the Kirin framework are based on blacklisting and must be defined upfront. This means that certain set of permissions would be considered as dangerous by the policy writer but any combination of these permissions that is not explicitly termed as dangerous is treated safe by default.

b) SCanDroid

Fuchs, et al. [23] proposed SCanDroid framework for Android to perform information flow analysis on applications in order to understand the flow of information from one component to another component. Consider a case where an application request permission to access multiple data stores i.e., public data store and private data store. The application requires permission for reading the data from the private store and writing data to the public store. SCanDriod analyzes the information flow of the application and report whether the application will transfer the information in the private store to the public store or not. However, SCanDroid also suffers from the same limitation of security policy expressibility. In order to consider some information flow to be dangerous, the

policy writers must define certain constraints prior to executing the policy. Similarly, if an information flow is not explicitly added to the set of constraints the framework will consider it to be safe.

c) Saint

Ongtang, et al. [13] proposed Saint – a framework that provides security policy constraints in a more expressive manner by defining install-time permission granting policies and runtime policies for inter-component communication. The framework places a number of dependency constraints on the permissions requested by the applications. These constraints may include the name of application, versions, signatures and set of other permissions. The effectiveness of Saint is based on its runtime policies for which reference monitors are used in the Android framework. The runtime policies are used to specify constraints for both the caller and callee applications. These constraints include permissions, configurations, signatures and/or the context in which the application is used e.g., time, location etc. The framework enables an application to protect and restrict its interface from being used by another application. However, this framework is not usercentric as it gives the option of policy specification to the application developers and not to the user. In our opinion, as the owner of the device, the decision to grant or deny access to the device resources should remain with the user and not with the application developers.

d) Apex

Apex [24] is an extension to the Android permission model that is more user-centric in allowing applications to access the device resources. Apex allows users to specify detailed runtime constraints to restrict the use of sensitive resources by applications. It is designed to overcome the limitation that the Android framework grants all the permissions to an application, which the application requests at install time. At install time the only way to deny the permissions requested by an application is to abort the installation. In the same way, the only way of revoking permissions once they are granted to an application is to uninstall the application. Contrary to this, Apex enables users to define conditions that must be fulfilled by an application in order to grant requested permissions to it. This means that it allows a subset of the requested permissions to be granted to the application at install-time. This way, user can start using the application with a limited number of permissions. The user may extend the granted permissions at a later stage. However, there are some limitations in the Apex framework. In the current Android architecture, the application developers assume that all the permissions that their application requests will be present in the manifest file. The developers often do not handle the unexpected security exceptions that are thrown when an application requests to access some resource(s) but the application does not have the required permissions to access it. If these exceptions are not properly handled – as may be the case in general – then we assume that most of the Android applications will not catch the exceptions and the exception will reach to the end of the call stack resulting in the termination of the thread.

e) Porscha

Ongtang, et al. [25] have proposed Porscha – a framework that enforces Digital Rights Management (DRM) policies – designed specifically for SMS, MMS and Email services allowing the content owners to restrict access to their content by specifying access control policies based on certain conditions like location and number of times to view a particular content etc. However, it is designed to facilitate different enterprises and government organizations with strictly controlled access policies.

f) CRePE

Conti, et al. [26] have proposed CRePE – a framework that enforces context-related fine-grained access control policies. It allows users to define policies that enable/ disable certain functionalities such as GPS, read SMS or Bluetooth discovery, based on the context of the device (e.g., location, noise, temperature, time, and nearby devices etc.). Furthermore, the context may also be defined by a trusted third party in scenarios where enterprise wide policy need to be deployed for all employees having Android smartphones. However, this framework only focuses on enabling/disabling of certain features of applications and cannot cope with the vulnerabilities that are formed by the permission usage across different applications.

g) XManDroid

Bugiel, et al. [27] have proposed XManDroid – extending monitoring on Android – to alleviate the problem of application-level privilege escalation attacks on Android. It analyzes the intercomponent communication mechanism among different applications in Android to ensure that these communication links comply with the predefined security policy. One of the major obstacles for the XManDroid is to define and maintain useful policies as well as policy exceptions. Some similar countermeasures against such malicious applications have already been proposed. Enck et al proposed "TaintDroid", a system-wide dynamic taint tracking system, in which multiple sources of sensitive data are tainted and the taint is used as a marker capable of real-time tracking of sensitive data [5]. They implemented "TaintDroid" and the evaluation results suggest that the overhead time for taint tracking is about 29% at most.

Takemori et al proposed the "white-list" measure which allows only secure and necessary applications to work [6]. In the white-list all approved

applications are shown. In order to prevent malicious applications from intruding, any unlisted application will be immediately deleted even if it is installed. Kawabata et al pointed out the risk that attackers could execute Java method by using JavaScript downloaded from the server and it may in turn cause malicious behavior [7]. To counter this, they proposed to conduct a static analysis for Android applications with JavaScript to ascertain its threat level.

Chin et al mentioned vulnerability of interapplication Communication in Android [8]. They pointed out that some malicious components can eavesdrop and tamper with the "Intent" while sending and receiving a message between applications. They surveyed 100 applications and revealed that they undoubtedly have such vulnerabilities. To realize this security goal without adding unnecessary burden to developers, Harunobu Agematsu proposed to prepare a dedicated API called "ADMS API" and to create "Knowledge Database" which security manager could use to judge malicious behavior [4].

The United States National Security Agency has recently announced the commencement of the SEAndroid (Security Enhanced Android) project as an addition to the Android kernel [20]. Similar to the well known and widely deployed SELinux Linux kernel patch, the SEAndroid project aims to establish a fine grained Mandatory Access Control model. It is further adjusted and extended to meet the requirements which arise on the Android platform, e.g., to secure inter process communication [21]. Once integrated into Android, SEAndroid may indeed prevent some of the attacks presented. SEAndroid is still in a very early development stage. It is unknown when or if SEAndroid will be integrated into the default code base of the operating system.

V. Summary and Conclusion

We have elaborated the limitations in the current Android security model in detail and in previous section we have presented the existing research proposals for improving overall Android security model. In this section we detail some of the security requirements that need to be taken into consideration while designing security mechanisms for smartphones in general and Android in particular. In order to alleviate the limitations and further strengthen the Android security model, one of the most important security concerns for the current smartphones is the lack of a model that allow users to specify, at a fine-grained level, which of the phone's resources should be accessible to third party applications. To design policies that are fine-grained in expressibility and are targeted to cater application-specific requirements is one of the biggest challenges in proposing new security enhancements. It requires a pre-design analysis of real applications to gather a larger collection of likely scenarios where the fine-grained policies are applicable. While designing a new framework, it should be capable of specifying a set of detailed runtime constraints to restrict the use of sensitive resources by applications. For example, the user may want to restrict certain applications to access a particular resource in a particular context (e.g., time, location, and maximum number of usage etc.) without uninstalling the application. This could be achieved by designing a framework that allows granting selective permissions at install time as well as monitor the use of these permissions at runtime by employing certain usage control mechanisms. The growing number of malware vulnerabilities has augmented serious concerns over security models for the smartphones. The recent attacks discussed in [17], [19], [20] have shown how easily some of the Android security features can be overturned by the malware developers. While designing new framework or proposing enhancement to the Android security model, we should consider that the model should not be by-passable by the sophisticated malware and/or the applications installed on the device as well as new applications. The framework should be designed in such a manner that it can validate that the system is not tampered with. It should be able to prevent information leakage from the device in scenarios where a legitimate application is replaced with a similar one containing Trojans that spy on user's sensitive information such as location, or, logs the phone calls and transfers that information to a remote server.

References Références Referencias

1. Google bets on Android future. http://news.bbc.co.uk/2/hi/technology/7266201.stm
2. Hongwei Luo†‡, Guili He†, Xiaodong Lin§, and Xuemin(Sherman) Shen‡ "Towards Hierarchical Security Framework for Smartphones" First IEEE International Conference on Communications in China: Communications Theory and Security (CTS) 2012.
3. P. Ferrill, "Exploring the android api," *Pro Android Python with SL4A*, pp. 113–138, 2011.
4. Harunobu Agematsu, Junya Kani, Kohei Nasaka, Hideaki Kawabata "A proposal to realize the provision of secure Android applications" IEEE 2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing.
5. J. Jamaluddin, N. Zotou and P. Coulton, "Mobile phone vulnerabilities: a new generation of malware," in *Consumer Electronics, 2004 IEEE International Symposium on*. IEEE, 2004, pp. 199–202.
6. William Enck, Peter Gilbert, Byung-Gon Chun, Landon P. Cox, Jaeyeon Jung, Patrick McDaniel, Anmol N. Sheth : "TaintDroid: An Information- Flow Tracking System for Realtime Privacy Monitoring on Smartphones", Proceedings of the 9th USENIX

Symposium on Operating Systems Design and implementation (OSDI'10), Canada, 2010.
7. Keisuke Takemori, Hideaki Kawabata, Takamasa Isohara, Ayumu Kubota, Jyunichi Ikeno: "Restriction framework for Android application -Whitelist-based installation", IPSJ (The Information Processing Society of Japan) SIG (The Special Interest Group) technical reports, 2011-CSEC-53-2, pp.1-6, 2011.5 (in Japanese).
8. Hideaki Kawabata, Takamasa Isohara, Keisuke Takemori, Ayumu Kubota: "Threat of Script abuse Android Permissions and Static Analysis", IPSJ SIG technical reports, 2011-CSEC-53-3, pp.1-6, 2011.5 (in Japanese).
9. A. Felt, M. Finifter, E. Chin, S. Hanna, and D. Wagner, "A survey of mobile malware in the wild," in *Proceedings of the 1st ACM workshopon Security and privacy in smartphones and mobile devices*. ACM, 2011, pp. 3–14.
10. Ruben Jonathan Garcia Vargas "Security Controls for Android" 2012 IEEE Fourth International Conference on Computational Aspects of Social Networks (CASoN) 215.
11. R. Rogers, J. Lombardo, Z. Mednieks, and B. Meike, *Android application development: Programming with the Google SDK*. O'Reilly Media, Inc., 2009.
12. C. Guo, H. Wang, and W. Zhu, "Smart-phone attacks and defenses," in *HotNets III*, 2004.
13. C. Dagon, T. Martin, and T. Starner, "Mobile Phones as Computing Devices: the Viruses Are Coming," IEEE Pervasive Computing, vol. 3, no. 4, 2004, pp. 11–15.
14. M. Goadrich and M. Rogers, "Smart smartphone development: ios versus android," in *Proceedings of the 42nd ACM technical symposium on Computer science education*. ACM, 2011, pp. 607–612.
15. SELinux project, "SEAndroid", August 2012.
16. S. Salah, S. Abdulhak, H. Sug, D. Kang, and H. Lee, "Performance analysis of intrusion detection systems for smartphone security enhancements," in Mobile IT Convergence (ICMIC), 2011 International Conference on. IEEE, 2011, pp. 15–19.
17. M. Pelino, Predictions 2010: Enterprise Mobility Accelerates Again, Forrester, 2009.
18. Angel Alonso Parrizas, SANS Institute, "Securely deploying Android devices", September 2011.
19. Jeter, L. Mani, M, Reinschmid, T., University of Colorado, "SmartPhone Malware: The danger and protective strategies", 2011 August.[20] Android Open Source Project, Google, "Android Security Overview", 2011.
20. S. Smalley, "SE Android release." SELinux Mailing List, Mailing List Archives (marc.info), January 2012. ttp://marc.info/?l=selinux&m=132588456202123&w=2.
21. National Security Agency, "SEAndroid Project Page,"January012.http://selinuxproject.org/page/SEAndroid.
22. W. Enck, M. Ongtang, and P. McDaniel, "On lightweight mobile phone application certification," in Proceedings of the 16th ACM conference on Computer and Communications Security. ACM, 2009, pp. 235–245.
23. A. Fuchs, A. Chaudhuri, and J. Foster, "Scandroid: Automated security certification of android applications," Manuscript, Univ. of Maryland, http://www.cs.umd.edu/˜avik/projects/scandroidascaa.
24. M. Nauman, S. Khan, and X. Zhang, "Apex: Extending android permission model and enforcement with user-defined runtime constraints," in Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security. ACM, 2010, pp. 328–332.
25. M. Ongtang, K. Butler, and P. McDaniel, "Porscha: Policy oriented secure content handling in android," in Proceedings of the 26th Annual Computer Security Applications Conference. ACM, 2010, pp. 221–230.
26. M. Conti, V. Nguyen, and B. Crispo, "Crepe: Context-related policy enforcement for android," Information Security, pp. 331–345, 2011.
27. S. Bugiel, L. Davi, A. Dmitrienko, T. Fischer, and A. Sadeghi, "Xmandroid: A new android evolution to mitigate privilege escalation attacks," TR-2011-04, Technische Universit¨at Darmstadt, April. 2011., Tech. Rep.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Network, Web & Security
Volume 13 Issue 12 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Cyber Police: An Idea for Securing Cyber Space with Unique Identification

By Ziaur Rahman, Md. Baharul Islam & A. H. M. Saiful Islam
Daffodil International University, Bangladesh

Abstract - The advancement of cyber technology completely depends on how conveniently we use it. Many people deceived in different ways in our current cyber system. We can prosper ourselves through the utilization of this internet technology. However any country can improve their online security through improving cyber system. On the other hand it may cause precarious outcome if it is incorrectly handled by any unplanned administration. An appropriate mechanism can move forward our cyber world with a safer e-biosphere. The purpose of this paper is to propose an idea that will ensure security and justice in cyber world. This idea proposes to diminish all types of anarchy from the cyber space by ensuring authentic identification to every internet user; securing website browsing; preventing any kind of fraud as well as guarantee truth and justice in the online world.

Keywords : *cyber world, wo rld wide web, virtual justification, cyber security, sensor technology.*

GJCST-E Classification : *K.4.4*

Strictly as per the compliance and regulations of:

Cyber Police: An Idea for Securing Cyber Space with Unique Identification

Ziaur Rahman [α], Md. Baharul Islam [σ] & A. H. M. Saiful Islam [ρ]

Abstract - The advancement of cyber technology completely depends on how conveniently we use it. Many people deceived in different ways in our current cyber system. We can prosper ourselves through the utilization of this internet technology. However any country can improve their online security through improving cyber system. On the other hand it may cause precarious outcome if it is incorrectly handled by any unplanned administration. An appropriate mechanism can move forward our cyber world with a safer e-biosphere. The purpose of this paper is to propose an idea that will ensure security and justice in cyber world. This idea proposes to diminish all types of anarchy from the cyber space by ensuring authentic identification to every internet user; securing website browsing; preventing any kind of fraud as well as guarantee truth and justice in the online world.

Keywords : cyber world, world wide web, virtual justification, cyber security, sensor technology.

I. Introduction

Initially we have tried to identify the reasons behind the existing cyber world have been working for long time since it was born. And finally we have found that the today's cyber space is unplanned, unmanaged, inconsistently distributed, commercialized, and immorally active. The relationship between the various attack methods and their corresponding solution are described (Adeyinka, 2008). Really there is no master plan under the huge cyber world. Managing a simple house is almost impossible without at least a minimum plan whereas this cyber world is running without of any reliable control.

For example, if a new application is developed, it is added to internet world spontaneously through enormous advertisement typically without any minimum certification or any valid justification as prerequisite. Consider an application for measuring the amount of love between two persons in percentage if just two names are given to it. A vast number of Apps and mini tools like this are unnecessarily available in different social network sites today. We have been using these or we have to insensibly click on it before we go on for further browsing without any evaluation even in some cases entirely unknowing what it is. We never consider how much time it wastes and how many problems it creates. Not only apps, we're unintentionally offered different irrelevant virus, worms etc through firing unwanted script when we download software or anything from different websites. Some websites and even many ads are always misusing our cookies and session to sneak our data irrespective of minimum privacy policy. We're technically deprived of doing anything against this type of deceptions we often face everywhere online as there is no exact authority to take care our allegations and to dispute it into a solution as well. However, up to now no endeavor we have towards establishing any regulatory body to step up against this kind of frauds. Internet Technology is now a matured youth after its disorderly long-teenage stage. But the bitter truth is that the amount of high speed internet users is not as enough as we ever hope. Whenever, this rate is quite alarming in developing countries around the globe. Unlimited commercialization of this very technology is not a new feature at all, but it has intractably increased in an intolerable extend from the last decades of the previous century. And the ultimate situation is getting worse than before day by day. Direct and indirect advertisements are now a critical hindrance while we surf internet. Immoral activities in cyber space are another very frequent issue from its birth to at present has noticeably caused hundreds of thousands of sufferers here and there across the world. It is said that cyber space is as a negative stuff as even several educated parents are totally unlikely to let their children to use internet at their early age. United Nations International Children's Emergency Fund (UNICEF) has recently funded a campaign through their website to save the Children abuse online. Certainly, it's not acceptable in any manner in this modern on-screen date. So far, now is the time even though it's quite late to rethink about our cyber space to advance it towards a wonderful virtual world that we all once dreamed of. To successfully encounter this challenge, we propose an awesome solution. Let's see what exactly Cyber Police wants to do. (Langner, 2013) finds out why cyber security risk cannot simply be "Managed" away. Towards measure internet security strategy for small medium enterprise is finding (Fortinet, 2013)] with better definition (Aspnes, 2003). (Library, 2009) worked with an annotated bibliography of select foreign-language academic literature. There were a lot of research on cyber crime study and their cost (Anderson, 2012).

Author α : Lecturer, Department of ICT, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh.
E-mail : zia@iut-dhaka.edu
Author σ : Senior Lecturer, Department of Multimedia Technology and Creative Arts, Daffodil International University, Dhaka, Bangladesh.
E-mail : baharul@daffidilvarsity.edu.bd
Author ρ : Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh.
E-mail : saifcse@daffodilvarsity.edu.bd

Cyber-crime is the big challenges for coming of age (David, 2012). Policy for cyber security is increasing via internet governance, reporting and awareness reforms (Amor, 2012).

II. Methodology

Diminishing the distance between virtual and the real world will be the preliminary task of Cyber Police. For this, first of all, we'll have to ensure a genuine online identity or the Cyber ID (CD) in short. It is really inevitable to have a unique online address as we usually have in our physical world. To guarantee this desired CD we can apply different modern way but now we want to begin with ratifying our email addresses. The email profile of a user exclusively has a duty to match his national citizen profile. To make it easy, we can regard national ID as our email address. Suppose, National id of a Bangladeshi traffic is 8217318598379 (as shown in the Figure 1) should be his distinctive CD. One of the prominent advantages of this system is that an infant will naturally be a member of this huge Cyber family immediate after its birth. However, almost half of the total idea will be impulsively implemented itself if we can make certain an authentic, credible and reliable CD.

Figure 1 : Sample National ID card of a Bangladeshi Citizen

Then the next steps will be easier to initiate as it requires not moving away much from the existing system structure currently we have. In our real life we habitually see, who you are entering into my house you must inform me before you go in. Same as, a user will have to substantiate his CD before logon into any web server or website. The server/site authority will let the specific traffic to sign in if no prior allegation is found alongside him. At the same time, it will also authenticate whether it's he or not exactly who is requesting for to login besides reporting his mental health as well. Question is how we can prove that the id and the user on behalf of it are uniquely resembles. The answer is so easy as we need not to be worried any kind. Modern finger print sensor technology is that what we can easily apply to solve this. In a broad sense if we consider it then we can think about DNA code encryption matching through a biometric encoding method. For these we only need an extra key on the keyboard or totally sensing input device. To check user's healthiness we'll just append a bio-informatics technology inside the additional desired button. For more reliability in future we can re-deign our system through making it artificially intelligent or through enabling its total object detection capability. Nonetheless, what about resource and time? Yes, No need to be nervous about it because we could save our valuable time and resource if the idea here is put into service completely. An activity history profile will simultaneously monitor and store user's event information of every second at server memory when he sends http request to the server. If he attempts to do anything out of security policy automatically will be logged out after informing the reason behind it. And it will also send a report (diary) to ICCJ accusing of the suspected user. International Cyber Court of Justice (ICCJ) is a proposed sovereign council to ensure cyber law and justice worldwide. ICCJ will receive and verify the diary immediate after submitting it and will go forward for further investigation to finally Figure out whether the suspect is guilty or not. If the acquisition is proved then the authority will take necessary step against the alleged user. As a penalty the accused id could be sent to the custody cell enlisting it as an illicit id for a certain period of time or a fixed balance could be cut down from his bank account as charge according to the intensity of the crimes. If ICCJ finds any severe crimes, it will redirect the case to the proper authority of the user's Country to warrant physical punishment for what he has ever done. Meanwhile a user will be privileged to appeal to ICCJ showing documents in favor of it to be acquitted himself. One may have a question how an unschooled inexpert user will do that. Yes, we can imagine e-lawyer to resolve this type of situation if it arises. The lawyer or anybody experienced will show the necessary browser record on behalf of the supposed user. If any traffic wants to convict against a website or server he could apply the same process above.

III. Experimental Result

To Login into the ICCJ email system we need to enter valid Cyber Id and password. If the username and password is correct then the traffic can sign in. Otherwise if the traffic enters wrong password or Cyber Id that doesn't match with the Cyber Id and Password exist in database he will be the shown the same interface again as we can see in Figure 2.

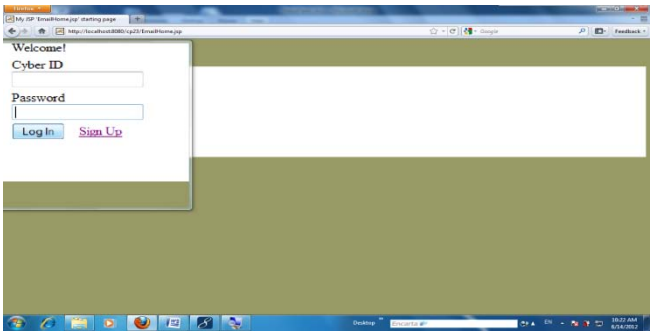

Figure 2 : Webmail login interface

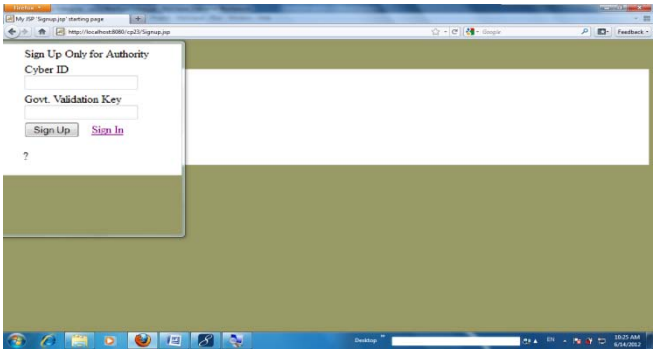

Figure 3 : Signup Interface System

In login page there is another option called signup out there for new entry beside form submit button. This idea demands a much authenticated type of login. As we said our national id will be our email id. And an email id will be generated by the government authority instead of individual initiation. But as here it's not very easy to implement this plan so we've considered a government validation key for temporarily entry to check up. For the real case or for professional use of this idea certainly we need our sophisticated cyber id (CD). To register a new account we need to click on Signup button. After clicking on it we see the windows as shown in Figure 3.

a) Sample Page Login Interface

After using his email a user can easily log out from the system. Now we will see how our ICCJ would be monitoring our internet surfing. And how will it control itself. First of all if we want to browse a website, as early as I enter the desired web address on our address bar; it will automatically redirect server login page as shown in the Figure 3 in the right interface. After login we will see our desired website in the left interface of the Figure 3 available below.

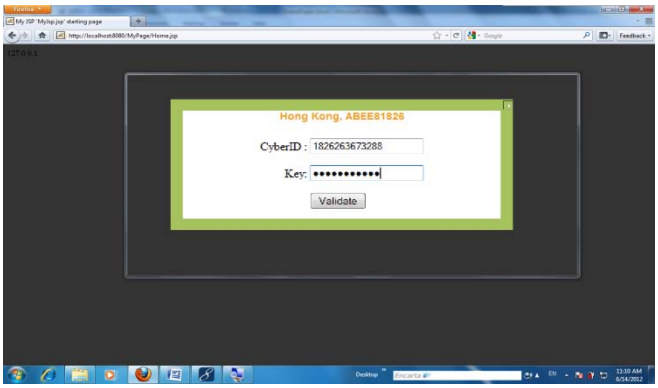

Figure 4 : Login window to access any website

Figure 5 : Ministry of Education in Bangladesh Website

After successfully checking up the validation, the user will see the page he looks for. Here as a sample we show our page. The page is official website of Ministry of Education, Bangladesh in Figure 5. In this page a webmail service is available for office use only. The webmail system is not public as it demands user authentication. If anybody unauthorized wants to use this webmail system and s/he enters wrong password and username more than three times then s/he will be logged out from that site or server. At the same time ICCJ will send an email to his email profile as an accused allegation.

b) Alleged User Email Interface

Once allegedly logging out from the server the user can't login into the same page again because of acquisition. He is supposed to open his ICCJ email inbox to check the allegation against him. Here in Figure 6 shows a sample message from ICCJ while a user is acquitted of. After reading the mail thrown by the ICCJ automated system if the user want to appeal in favor of his own side he can do it before the deadline. For the successful implementation of this proposed idea we certainly need an active browser what can provide us necessary history on browsing period. The message will

show an attachment containing the activities details what he did online that is shown in Figure 7.

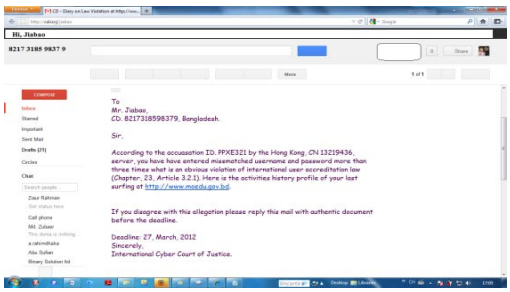

Figure 6 : Email from ICCJ immediate after allegation

Activities History Profile

CD. 8217318598379

Server : Hong Kong. CN13219436

No	Even ID	Time	Purpose	Remark
1	AB9871	16:34:26	Redirecting Server Validation	√
2	EE2144	16:34:26	Pressing Sensor Device	√
3	BE2513	16:34:27	Clicking on Validation Button	√
4	AE7632	16:34:28	Redirecting www.moedu.gov.bd	√
5	CF2653	16:34:30	Clicking on Webmail Button	√
6	EC6533	16:34:33	Submitting username and password	x
7	DE1231	16:34:35	Submitting username and password	x
8	EE6723	16:34:36	Submitting username and password	x
9	DD5421	16:34:36	Logging out by server	√

Figure 7 : Active history profile

c) Suspending (Punishment) Message from ICCJ

Finally if the allegation is proved the ICCJ authority will suspend his Cyber ID for a certain period of time as shown in Figure 8.

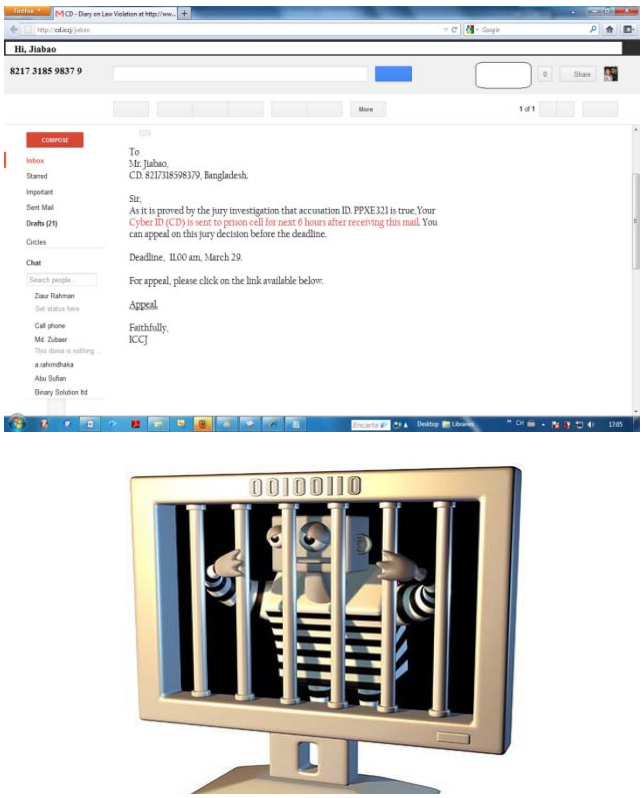

Figure 8 : Message from ICCJ containing suspension report and cyber ID in the prison

Within this time period the user will not be able to surf internet. If he tries to do the same he will be given this message as shown in Figure 9.

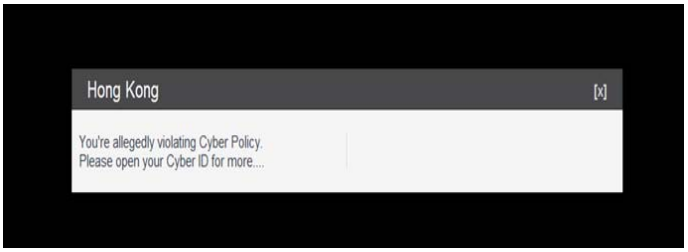

Figure 9 : User Window While CD is allegedly accused of

After the deadline the cyber id will be freed (as it was in the prison database Figure 9 specified right) and it will work again as it did before.

IV. Discussion

Our idea is successfully tested with the help of certain necessary tools. Before testing there were some faults and bugs inside it but now it is totally bug free. The schedule of test is given in table 1.

Table 1 : Test Schedule

Test Start Date	Test Complete Date	What was Tested	What was not Tested
June 1, 2013	June 12, 2012	• Login Security • Hosting • Deployment • Filtering • Session • Cookie	• Finger Print Validation • National ID Matching

V. Conclusion

We know what is happening in online world. Online technology is a wonderful invention for the humankind but we're not able to get maximum number of throughput from it because of uncontrolled and so called regulation system. Now it is the time for change. Only a convenient change can be a solution for the problems we have ever encountered. So far, though it's quite late to rethink about our cyber space to advance it towards a wonderful virtual world as once our anterior generation dreamed of. To successfully encounter this challenge "Cyber Police: An Idea" we propose may be an awesome solution. It is true that founding and activating an international council like ICCJ is the foremost important task before apply this initiative titling "Cyber Police: An Idea" and it's not as easy as we're talking about. But to save our cyber space for a better, safer virtual world the United Nations or the leading IT Giants today can play a vital role to implement this dream towards true. We keep our hope alive.

References Références Referencias

1. Adeyinka, O. (2008) 'Internet Attack Methods and Internet Security Technology', *Second Asia International Conference on Modelling & Simulation,* 77-82.
2. Amor, F.E.B. (2012), 'Policy Memorandum: Increase Cyber-Security via Internet Governance, Reporting and Awareness Reforms, The Journal of Science Policy and Governance, 2(1).
3. Anderson, et al. (2012), 'Measuring the Cost of Cyber crime, 1-31.
4. Aseef, N. et al (2005), 'Cyber-Criminal Activity and Analysis', White Paper, Center for Security & Privacy Solutions.
5. Aspnes, J., Feigenbaum, J., Mitzenmacher, M., Parkes, D., (2003) 'Towards Better Definitions and Measures of Internet Security', Position Paper.
6. Cleveland, F.M. 2008, Cyber Security Issues for Advanced Metering Infrastructure (AMI), IEEE.
7. Comprehensive Study on Cybercrime, UN Office on Drugs and Crime, February 2013.
8. David H. (2010), 'Cybercrime Coming of Age', White paper.
9. Doyle, C., (2013), 'Cyber-security: Cyber Crime Protection Security Act (S. 2111, 112th Congress)-A Legal Analysis, CRS Report for Congress Research Service.
10. Ericsson, G.N., 2010, 'Cyber security and power system communication-Essential parts of smart grid infrastructure', IEEE Transactions on Power Delivery, 25 (3).
11. François P. (2010) Cybercrime and Hacktivism, White Paper, McAfee Labs.
12. Fortinet, (2013) 'Cyber criminals Today Mirror Legitimate Business Processes', Cyber crime Report.
13. Grzybowski, K. M., (2012), 'An Examination of Cyber-crime and Cyber-crime Research: Self-control and Routine Activity Theory, Arizona State University.
14. Guerra, P., (2009) How Economics and Information Security Affects Cyber Crime and What It Means in the Context of a Global Recession, White paper.
15. Langner, R., Pederson, P., (2013) 'Bound to Fail: Why Cyber Security Risk Cannot Simply Be "Managed" Away', Cyber Security series, Foreign Policy at Brookings.
16. Library, C. (2009), 'Cyber-crime: An annotated bibliography of select foreign-language academic literature', Federal Research Division, Library of Congress.
17. Library, P. (2011), 'Cyber crime: Issues (Background Paper), Library of Parliament, Ottawa, Canada, Publication No. 2011-36-E.
18. Ponemon, I., (2011), Second Annual Cost of Cyber Crime Study, Benchmark Study of U.S. Companies, Research Report.
19. Software, G.F.I. (2011) 'Towards a comprehensive Internet security strategy for SMEs', GFI White Paper, 1-6.
20. Schaeffer, B.S., Chan, H., Ogulnick, S., (2009), 'Cyber Crime and Cyber Security', White paper.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Network, Web & Security
Volume 13 Issue 12 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

An Enhanced Web Data Learning Method for Integrating Item, Tag and Value for Mining Web Contents

By R. Marutha Veni & P. Kavipriya
CMS College of Science and Commerce, India

Abstract - The Proposed System Analyses the scopes introduced by Web 2.0 and collaborative tagging systems, several challenges have to be addressed too, notably, the problem of information overload. Recommender systems are among the most successful approaches for increasing the level of relevant content over the "noise." Traditional recommender systems fail to address the requirements presented in collaborative tagging systems. This paper considers the problem of item recommendation in collaborative tagging systems. It is proposed to model data from collaborative tagging systems with three-mode tensors, in order to capture the three-way correlations between users, tags, and items. By applying multiway analysis, latent correlations are revealed, which help to improve the quality of recommendations. Moreover, a hybrid scheme is proposed that additionally considers content-based information that is extracted from items.

We propose an advanced data mining method using SVD that combines both tag and value similarity, item and user preference. SVD automatically extracts data from query result pages by first identifying and segmenting the query result records in the query result pages and then aligning the segmented query result records into a table, in which the data values from the same attribute are put into the same column. Specifically, we propose new techniques to handle the case when the query result records based on user preferences, which may be due to the presence of auxiliary information, such as a comment, recommendation or advertisement, and for handling any nested-structure that may exist in the query result records.

GJCST-E Classification : H.2.8

Strictly as per the compliance and regulations of:

An Enhanced Web Data Learning Method for Integrating Item, Tag and Value for Mining Web Contents

R. Marutha Veni [α] & P. Kavipriya [σ]

Abstract - The Proposed System Analyses the scopes introduced by Web 2.0 and collaborative tagging systems, several challenges have to be addressed too, notably, the problem of information overload. Recommender systems are among the most successful approaches for increasing the level of relevant content over the "noise." Traditional recommender systems fail to address the requirements presented in collaborative tagging systems. This paper considers the problem of item recommendation in collaborative tagging systems. It is proposed to model data from collaborative tagging systems with three-mode tensors, in order to capture the three-way correlations between users, tags, and items. By applying multiway analysis, latent correlations are revealed, which help to improve the quality of recommendations. Moreover, a hybrid scheme is proposed that additionally considers content-based information that is extracted from items.

We propose an advanced data mining method using SVD that combines both tag and value similarity, item and user preference. SVD automatically extracts data from query result pages by first identifying and segmenting the query result records in the query result pages and then aligning the segmented query result records into a table, in which the data values from the same attribute are put into the same column. Specifically, we propose new techniques to handle the case when the query result records based on user preferences, which may be due to the presence of auxiliary information, such as a comment, recommendation or advertisement, and for handling any nested-structure that may exist in the query result records.

I. Introduction

a) Data Mining

Data mining (the analysis step of the "Knowledge Discovery in Databases" process, or KDD), is a field at the intersection of computer science and statistics, is the process that attempts to discover patterns in large data sets. It utilizes methods at the intersection of artificial intelligence, machine learning, statistics, and database systems. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data preprocessing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

The term is a buzzword, and is frequently misused to mean any form of large-scale data or information processing (collection, extraction, warehousing, analysis, and statistics) but is also generalized to any kind of computer decision support system, including artificial intelligence, machine learning, and business intelligence. In the proper use of the word, the key term is discovery, commonly defined as "detecting something new". Even the popular book "Data mining: Practical machine learning tools and techniques with Java" (which covers mostly machine learning material) was originally to be named just "Practical machine learning", and the term "data mining" was only added for marketing reasons. Often the more general terms "(large scale) data analysis", or "analytics" – or when referring to actual methods, artificial intelligence and machine learning – are more appropriate. According to one source, data mining is a marketing term coined by HNC, a San Diego-based company (now merged into FICO), at the beginning of the century to pitch their Data Mining Workstation.

The actual data mining task is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection) and dependencies (association rule mining). This usually involves using database techniques such as spatial indexes. These patterns can then be seen as a kind of summary of the input data, and may be used in further analysis or, for example, in machine learning and predictive analytics. For example, the data mining step might identify multiple groups in the data, which can then be used to obtain more accurate prediction results by a decision support system. Neither the data collection, data preparation, nor result interpretation and reporting are part of the data mining step, but do belong to the overall KDD process as additional steps.

b) Web Mining

i. Web Structure Mining

Web structure mining is the process of using graph theory to analyze the node and connection structure of a web site. According to the type of web

Author α : MCA, Mphil, Asst. Professor, Dept of Computer Science, Dr. SNS RCAS, Coimbatore. E-mail : dhanuvene1@gmail.com
Author σ : MCA,(Mphil), Bharathiar University, Asst. Professor, School of commerce, CMS College of science and commerce, Coimbatore. E-mail : kavipriya.rajen@gmail.com

structural data, web structure mining can be divided into two kinds:

1. Extracting patterns from hyperlinks in the web: a hyperlink is a structural component that connects the web page to a different location.
2. Mining the document structure: analysis of the tree-like structure of page structures to describe HTML or XML tag usage.

ii. *Web Content Mining*

Mining, extraction and integration of useful data, information and knowledge from Web page contents. The heterogeneity and the lack of structure that permeates much of the ever expanding information sources on the World Wide Web, such as hypertext documents, makes automated discovery, organization, and search and indexing tools of the Internet and the World Wide Web such as Lycos, Alta Vista, WebCrawler, ALIWEB, MetaCrawler, and others provide some comfort to users, but they do not generally provide structural information nor categorize, filter, or interpret documents. In recent years these factors have prompted researchers to develop more intelligent tools for information retrieval, such as intelligent web agents, as well as to extend database and data mining techniques to provide a higher level of organization for semi-structured data available on the web. The agent-based approach to web mining involves the development of sophisticated AI systems that can act autonomously or semi-autonomously on behalf of a particular user, to discover and organize web-based information.

II. Literature Review

a) Mining Web Session Characteristic for Boundary Defense based on Hidden Markov Model

Yi Xie* and Xiangnong Huang

i. *Proposal*

Different from most existing studies on Web session identification, a novel dynamic real time user session processes description method is presented in this paper. The proposed scheme doesn't rely on presupposed threshold or client/server side data which are widely used in traditional session detection approaches. A new parameter is defined based on inter-arrival time of HTTP requests. A nonlinear algorithm is introduced for quantization. Nonparametric hidden semi-Markov model is applied to distinguish the user session processes. A probability function is derived for predicting user session processes. Experiments based on real HTTP traces of large-scale Web proxies are implemented to valid the proposal.

ii. *Drawbacks of the Proposal*

Different from traditional user session model, in this paper, a user session of a special user is divided into three segments: activity period, silent period and off-lining period. Activity period means the user is surfing the Internet, which causes the frequent interactions between user and different remote servers. Silent period indicates network connection is enable, but the user does nothing. During this period, the HTTP requests are mainly launched by softwares instead of user's own action. Thus, the number of requests in this period is far less than that of activity period. The last period means the network connection is unworkable or user has left.

b) Applying Concept Analysis to User- Session- Based Testing of Web Applications

i. *Proposal*

The continuous use of the Web for daily operations by businesses, consumers, and the government has created a great demand for reliable Web applications. One promising approach to testing the functionality of Web applications leverages the user session data collected by Web servers. User-session-based testing automatically generates test cases based on real user profiles. The key contribution of this paper is the application of concept analysis for clustering user sessions and a set of heuristics for test case selection. Existing incremental concept analysis algorithms are exploited to avoid collecting and maintaining large user-session data sets and to thus provide scalability. We have completely automated the process from user session collection and test suite reduction through test case replay. Our incremental test suite update algorithm, coupled with our experimental study, indicates that concept analysis provides a promising means for incrementally updating reduced test suites in response to newly captured user sessions with little loss in fault detection capability and program coverage.

ii. *Drawbacks of Proposal*

One approach to testing the functionality of Web applications that addresses the problems of the path based approaches is to utilize capture and replay mechanisms to record user-induced events, gather and convert them into scripts, and replay them for testing, Tools such as Web King and Rational Robot provide automated testing of Web applications by collecting data from users through minimal configuration changes to the Web server. The recorded events are typically base requests and name-value pairs (for example, form field data) sent as requests to the Web server. A base request for a Web application is the request type and resource location without the associated data. To our knowledge, these techniques do not include incremental approaches to test suite reduction.

c) Clustering and Tailoring User Session Data for Testing Web Applications

i. *Proposal*

Web applications have become major driving forces for world business. Effective and efficient testing of evolving web applications is essential for providing reliable services. In this paper, we present a user

session based testing technique that clusters user sessions based on the service profile and selects a set of representative user sessions from each cluster. Then each selected user session is tailored by augmentation with additional requests to cover the dependence relationships between web pages. The created test suite not only can significantly reduce the size of the collected user sessions, but is also viable to exercise fault sensitive paths. We conducted two empirical studies to investigate the effectiveness of our approach one was in a controlled environment using seeded faults, and the other was conducted on an industrial system with real faults. The results demonstrate that our approach consistently detected the majority of the known faults by using a relatively small number of test cases in both studies.

ii. *Drawbacks of the Proposal*

User-session based testing makes use of field data to create test cases, which has the great potential to efficiently generate test cases that can effectively detect residual faults. However, this approach is relatively new compared to traditional well developed techniques. There are several issues that must be addressed before it can serve as a sole testing method in practice. For an application that has been in production for a long time, the number of user sessions can be extremely large. Using all of the collected user session data requires much effort to determine which portion of the data can serve as the best representative of the system behavior. Nevertheless, a vast number of user sessions may not necessarily guarantee good coverage of the expected system behavior.

d) Separating Interleaved User Sessions from Web Log

i. *Proposal*

Analysis of user behavior on the Web presupposes a reliable reconstruction of the users' navigational activities. The quality of reconstructed sessions affects the result of Web usage mining. This paper presents a new approach for interleaved server session from Web server logs using m-order Markov model combined with a competitive algorithm. The proposed approach has the ability to reconstruct interleaved sessions from server logs. This capability makes our work distinct from other session reconstruction methods. The experiments show that our approach provides a significant improvement in regarding interleaved sessions compared to the traditional methods.

ii. *Drawbacks of the Proposal*

Session reconstruction is an essential data pre-process step in Web usage mining. The primary session reconstruction approaches based on time and reference cannot reconstruct the interleaved sessions and perform poorly when the client's IP address is not available. In this paper, an algorithm based on m-order Markov model is proposed which can reconstruct interleaved sessions from Web logs. The experiments show the promising result that the m-order Markov model has the ability to divide interleaved sessions, and can provide a further improvement when it is combined with the competitive approach.

e) Optimal Algorithms for Generation of User Session Sequences Using Server Side Web User Logs

i. *Proposal*

Identification of user session boundaries is one of the most important processes in the web usage mining for predictive prefetching of user next request based on their navigation behavior. This paper presents new techniques to identify user session boundaries by considering IPaddress, browsing agent, intersession and intrasession timeouts, immediate link analysis between referred pages and backward reference analysis without searching the whole tree representing the server pages. A complete set of user session sequences and the learning graph based on these user session sequences is also generated. Using this graph predictive prefetching is done. Comparison on the performance of the given approach with the existing reference length method and maximal reference method was done. Our analysis with different server's logs shows that our approach provides better results in terms of time complexity and precision to identify user session boundaries and also to generate all the relevant user session sequences.

ii. *Drawbacks of Proposal*

The analysis indicated that the existing web technology faces so many problems. One among them is personalization of web pages. Personalization is achieved if we know the browsing pattern of users. Our algorithm generates the efficient user session sequences with the less time complexity and good accuracy compared to the existing works. Thus we can reduce the latency. In the forth-coming papers USIDALG can be modified to generate the efficient learning graph to predict and prefect the user's next request.

iii. *The Existing Researches*

The Existing researches in personalizing the web user were single entity based and a summary of few researches are presented and the proposed system is developed by clearly understanding the below problems. This chapter discusses some of the existing techniques presented by different authors.

1. R. Cooley, B. Mobasher, and J. Srivastava, proposed a system for Navigation, in contrast to search, generally requires hierarchical storage, i.e. users need to create folders or directories and to store the information items "inside" them in preparation for future retrieval and use. Although other navigation methods have been proposed such as faceted classification and hypertext, neither is in

common usage in widely used operating systems, so we restrict our discussion here to common hierarchical methods. Hierarchical storage was first introduced to end-users in the *Multics* operating system in the mid 60s. Users were allocated a personal directory, in which they could create their own subdirectories, sub-subdirectories, etc., and store their files in any of these "locations." This directory structure was later applied in the Unix and the Linux operating systems.

2. O. Nasraoui, R. Krishnapuram, and A. Joshi, worked on the location metaphor became even clearer with the creation of digital folders first introduced in the *Xerox Star* in 1981. A folder is a visual metaphor for a location: users can see information items "inside" folders, as well as manipulate items and folders in various straightforward ways, e.g. drag and drop information items from one folder to another, etc.1 This folder hierarchy metaphor was later applied by Apple in the Mac operating systems and then by Microsoft in their Windows operating systems. Thus, location-based storage has been used without significant modifications, continuously and almost exclusively for several decades.
3. O. Nasraoui, R. Krishnapuram, H. Frigui, and A. Joshi, proposed a survey of unsupervised and semi-supervised clustering methods was presented by Grira, Crucianu and Boujemaa in. Squared error algorithms rely on the possibility of representing each cluster by a prototype. In general, the prototypes are the cluster centroids, as in the Kmeans algorithm. Fuzzy versions of methods based on the squared error are also defined, such as the Fuzzy C-Means. When compared to their 'crisp' counterparts, fuzzy methods are more successful in avoiding local minima of the cost function and can model situations where clusters actually overlap.
4. J. Srivastava, R. Cooley, M. Deshpande, and P.-N. Tan In Morzy et al, Proposed a bottom-up approach of clustering based on Web Access Sequences is given, where frequent sequence patterns among web user sessions are identified. The users are then clustered based on their access sequence similarity. Shi has used the approach of fuzzy modelling taking into account the time duration that a user spends at a URL. Nasraoui et al have used the Competitive Agglomeration algorithm for Relational Data which yielded optimal number of clusters with non-Euclidean measures. In, it is argued that web user session identification itself is a non-trivial issue and clustering techniques have been used to characterise a user session. gives a basis of evaluating web usage mining approaches and for predicting the user's next request.
5. M. Spiliopoulou and L.C. Faulstich presented A survey of classification in data mining is given in. A sequence based clustering for web usage mining using K-means algorithm with artificial neural networks and Markov models is given in. It also demonstrates how a fuzzy approach yields superior accuracy. Artificial neural networks have been proven to be effective in dealing with classification problems and other machine learning areas. contains a brief tutorial of ANNs referred to in Section 6 and 7. Multilayered Perceptrons (MLP) were found to be appropriate for the dataset used. The applicability of MLPs is discussed. talks about Naive Bayes classifier which assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. Prefetching has been applied to a variety of distributed and parallel systems to hide communication latency
6. T. Yan, M. Jacobsen, H. Garcia-Molina, and U. Dayal, Crovella and Barford experimented and analyzed the effect of prefetching on the network performance by considering network delay as the primary cost factor. A simple transport rate controlled mechanism was proposed to improve the network performance. The usage of anchor text to index URL's in Google search engine was suggested by Brin and Page. The research focused on effective usage of additional information present in the hypertext. Chakrabarti et al designed and evaluated an automatic resource compilation system that could perform analysis of text and links to determine the web resources suitable for a particular topic.
7. M. Perkowitz and O. Etzioni Davison conducted a detailed analysis that focused on examining the descriptive quality of web pages and the presence of textual overlap in web pages. A keyword-based semantic prefetching approach was designed by Cheng and Ibrahim that could predict future requests based on semantic preferences of past retrieved Web documents. The scheme was evaluated by considering the Internet news services. Neural network was applied over the keyword set to predict future requests.
8. J. Borges and M. Levene, R.Cooley developed a quantitative model based on support logic that used information such as usage, content and structure to automatically identify interesting knowledge from web access patterns. The effectiveness of link-based and content-based ranking method in finding the web sites was analyzed by Craswell et al. The results indicated that anchor texts are highly useful in site finding.
9. O. Zaiane, M. Xin, and J. Han, Davison proposed a text analysis method that examined web page content to predict user's next request. The algorithm used text in and around the hypertext anchors of selected web pages to determine user's interest in

accessing web pages. Chen et al proposed a framework that used link analysis algorithm for exploiting both explicit (hyperlinks embedded in web page) and implicit (imagined by end-users) link structures. The framework had the ability to analyze interactions between users and the web.

10. O. Nasraoui and R. Krishnapuram developed a model including PageRank and HITS (Hypertext Induced Topic Selection) were the most popular webpage ranking algorithms. HITS emphasized on mutual reinforcement between the authority and hub web pages, whereas PageRank emphasized on hyperlink weight normalization. Ding et al generalized the concepts of mutual reinforcement and hyperlink weight normalization into a unified framework. Nadav and Kevin investigated anchor text to observe its relationship to titles, frequency of queries satisfied and the homogeneity of results obtained. Analysis indicated the anchor text resembled real-world queries in terms of its term distribution and length.
11. O. Nasraoui, C. Cardona, C. Rojas, and F. Gonzalez, Zhuge defined semantic links between resources to establish a high-level single semantic image to improve the quality of search result sets. The mathematical notations and formal structure of the semantic link network was presented in. Pierrakos et al clearly analyzed and presented the web usage mining process such as data collection, data preprocessing and pattern discovery that could be applied for web personalization. Jung carried out semantic outlier detection and segmentation using online web request streams to infer the relationships among web requests. A user support mechanism based on knowledge sharing with users through collaborative web browsing was proposed in [18]. It mainly focused on extracting user's interests from their own bookmarks.
12. P. Desikan and J. Srivastava, & Alexander Pons [19] proposed a technique that semantically bundled objects from slower loading web pages with objects of faster loading web pages. It was done to prefetch objects for the client's system prior to accessing the slower loading web page. carried out analysis on the web pages of different categories from Open Directory Project (ODP). They suggested the use of cohesive and non-cohesive text present near the anchor text to extract information about the target web page.
13. O. Nasraoui, C. Rojas, and C. Cardona researched a Access sequences as a criterion is not primary because these can be misleading in cases where the user does not know the ideal route to his destination. Also, considering sequences by themselves as a parameter has the risk of incorporating the undesirable step of giving equal importance to all sites, irrespective of the amount of time spent there, due to which the focus of the analysis is lost. In this paper, the time spent by a user at a URL is the criterion for analysing his degree of interest. The Naive Bayes Classifier is applied, following which, the K-means classification algorithm (statistical) is then compared with the Multilayer Perceptron (artificial neural networks) method using logged web usage data to analyse accuracy in classification.
14. C. Burges, T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender, proposed Search as an Alternative to Navigation Through most of its long history, the hierarchical method has met with criticism. One disadvantage is that classification of information can 'hide' it from the user, and therefore reduce the chances of quick retrieval or reminding. In addition, the act of categorisation is itself cognitively challenging; users may find it hard to categorise information that could be stored in more than one category. Categorisation is also difficult because it requires that people anticipate future usage; moreover, that usage may change over time. At retrieval time, users need to recall how information was classified, which can be difficult when there are multiple categorisation possibilities. These problems were illustrated in a study of email categorisation. They found that users with many categories found it harder to file, and were more likely to create spurious unused folders. These apparent problems with navigation caused many PIM researchers and software developers to turn to *Search* as an alternative. There are intuitive potential advantages of search for both retrieval and organization. Search promises to be more flexible and efficient at *retrieval*, it does not depend on remembering the correct storage location; instead, users can specify in their query any attribute they happen to remember. They can also retrieve information via a single query instead of using multiple operations to laboriously navigate to the relevant part of their folder hierarchy. Regarding storage, search potentially finesses the *organizational problem* - as users don't have to engage in complex organizational strategies that exhaustively anticipate their future retrieval requirements. These arguments against navigation have been bolstered by recent developments in web access, where initial use of navigational systems such as Yahoo.
15. K. W. Church, W. Gale, P. Hanks, and D. Hindle, superseded a search engines such as Google The same logic has led to the development of experimental PIM search engines such as *Phlat*, *SIS*, *Haystack*, and *Raton Laveur*, as well as commercial systems such as *Einfish Personal*, *Copernic Desktop Search*, *Yahoo! Desktop Search* and *Microsoft Desktop Search*. Some more radical

systems such as *Lifestreams*, *Canon Cat, Presto*, *Placeless Documents*, *MyLifeBits*, and *Swiftware* explore alternatives to location-based hierarchies. However despite the rapid development of new such technologies, we know little about the effects of improved desktop search on user behaviour. In this study therefore we set out to test the following predictions about the effects of desktop search on both file retrieval and organization:

Retrieval: Search is more efficient and flexible for retrieval, thus improved quality of search engines should lead to a substantial increase in file search and eventually a preference for search over navigation.
File Organization: Users are known to have problems organizing files effectively for retrieval. Search allows retrieval without such manual organization and improved search should lead to a reduced use of filing strategies in preparation for later retrieval.

16. Q. Gan, J. Attenberg, A. Markowetz, and T. Suel worked on a Navigation or Search: Prior Evidence Pertaining to the Debate Evidence concerning users' *search* preference comes from empirical studies that examine retrieval behaviour. An early paper concerning users' retrieval habits, combined Barreau's interviews of novice personal computer users (using DOS, Windows 3.1 and OS/2) with Nardi's interviews of experienced Macintosh users. In both cases, users "overwhelmingly" preferred to navigate to their files than to search for them. Similar preferences for navigation were obtained in other more recent studies. These early findings raise a question– if search better suits users' requirements, why do they prefer navigation? One argument is that search technology is still immature. For example, Fertig and his colleagues argued that these navigation preferences result from limitations in search technology, and that improvements in search would inevitably lead to the replacement of navigation. They noted that the PIM search engines of that time (the mid 90s) were "slow, difficult, or only operate on file names (not content)" and did not provide incremental indexing. Fertig et al. further speculated that "inclusion of these better search techniques into current systems could sway results". However, their claim that the improvement of search engines would lead to an increased preference for search over navigation has not been tested empirically.
17. T. Joachims, presented few Other evidence challenging the effects of improved *search* concerns users' organizational efforts to prepare for future retrieval. There is some evidence that users seem to want to preserve folders, even when improved search is possible. Jones, Phuwanartnurak, Gill, & Bruce asked [14] participants the following question: "Suppose you could find your personal information using a simple search rather than your current folders.… Can we take your folders away?" Only one participant responded positively. In contrast, Dumais et al.'s participants tended to mildly agree with the sentence "I would likely to put less effort into maintaining a detailed set of folders for my files if I could depend on SIS (i.e., the *Stuff I've Seen* search engine) to find what I am looking for". Both studies asked whether the use of improved search engines would lead to less reliance on folders, but (perhaps because Jones et al. asked the question in a more extreme way) received different answers. Notice, however, that both researchers asked this as a hypothetical question.
18. K. W.-T. Leung, W. Ng, and D. L. Lee proposed lot of Improvements in Desktop Search Engines Today, more than a decade after Fertig et al.'s claims, commercial PIM search engines have improved considerably, newer search engines (such as *Google Desktop* and *Spotlight*) are better than the older ones (such as *Windows XP Search Companion* and Mac *Sherlock*) in the following ways:

Cross-format search: One limit of older search engines was that they allowed users to search only one format at a time. Following the SIS initiative, several improved search engines now support search across multiple datatypes – files, emails, instant messages and Web history within the same search query. This allows them to address the project fragmentation problem, where information items related to the same project but in different formats, are stored in different locations.

f) Existing System

The three types of recommendations in STSs (i.e., item, tag, and user recommendations) have been so far addressed separately by various approaches, which differ significantly to each other and have, in general, an ad hoc nature. Since in STSs all three types of recommendations are important, what is missing is a unified framework that can provide all recommendation types with a single method. Moreover, existing algorithms do not consider the three dimensions of the problem. In contrast, they split the threedimensional space into pair relations {user, item}, {user, tag}, and {tag, item}, that are two-dimensional, in order to apply already existing techniques like CF, link mining, etc. Therefore, they miss a part of the total interaction between the three dimensions. What is required is a method that is able to capture the three dimensions all together without reducing them into lower dimensions.

Finally, the existing approaches fail to reveal the latent associations between tags, users, and items. Latent associations exist due to three reasons:

1. Users have different interests for an item,
2. Items have multiple facets, and

3. Tags have different meanings for different users.

As an example, assume two users in an STSs for Web bookmarks (e.g., Del.icio.us, Bibsonomy). The first user is a car fan and tags a site about cars, whereas the other tags a site about wild cats. Both use the tag "jaguar." When they provide the tag "jaguar" to retrieve relevant sites, they will receive both sites (cars and wild cats). Therefore, what is required is a method that can discover the semantics that are carried by such latent associations, which in the previous example can help to understand the different meanings of the tag "jaguar."

III. Methodology

a) Singular Value Decomposition

Let X denote an m x n matrix of real-valued data and *rank* r, where without loss of generality $m \geq n$, and therefore $r \leq n$. In the case of microarray data, x_{ij} is the expression level of the i^{th} gene in the j^{th} assay. The elements of the i^{th} row of X form the n-dimensional vector $\boldsymbol{g}_i$, which we refer to as the *transcriptional response* of the i^{th} gene. Alternatively, the elements of the j^{th} column of X form the m-dimensional vector $\mathbf{a}_j$, which we refer to as the *expression profile* of the j^{th} assay.

The equation for singular value decomposition of X is the following:

$$X = USV^{\mathrm{T}} \tag{5.1}$$

Where U is an m x n matrix, S is an n x n diagonal matrix, and V^T is also an n x n matrix. The columns of U are called the *left singular vectors*, $\{\mathbf{u}_k\}$, and form an orthonormal basis for the assay expression profiles, so that $\mathbf{u}_i \cdot \mathbf{u}_j = 1$ for $i = j$, and $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ otherwise. The rows of V^T contain the elements of the *right singular vectors*, $\{\mathbf{v}_k\}$, and form an orthonormal basis for the gene transcriptional responses. The elements of S are only nonzero on the diagonal, and are called the *singular values*. Thus, $S = \mathrm{diag}(s_1,\ldots,s_n)$. Furthermore, $s_k > 0$ for $1 \leq k \leq r$, and $s_i = 0$ for $(r+1) \leq k \leq n$. By convention, the ordering of the singular vectors is determined by high-to-low sorting of singular values, with the highest singular value in the upper left index of the S matrix. Note that for a square, symmetric matrix X, singular value decomposition is equivalent to diagonalization, or solution of the eigenvalue problem.

One important result of the SVD of X is that

$$X^{(l)} = \sum_{k=1}^{l} \mathbf{u}_k s_k \mathbf{v}_k^{\mathrm{T}} \tag{5.2}$$

is the closest rank-l matrix to X. The term "closest" means that $X^{(l)}$ minimizes the sum of the squares of the difference of the elements of X and $X^{(l)}$, $\sum_{ij} | x_{ij} - x^{(l)}_{ij} |^2$.

One way to calculate the SVD is to first calculate V^T and S by diagonalizing X^TX:

$$X^{\mathrm{T}} X = VS^2V^{\mathrm{T}} \tag{5.3}$$

and then to calculate U as follows:

$$U = XVS^{-1} \tag{5.4}$$

where the $(r+1),\ldots,n$ columns of V for which $s_k = 0$ are ignored in the matrix multiplication of Equation 5.4. Choices for the remaining n-r singular vectors in V or U may be calculated using the Gram-Schmidt orthogonalization process or some other extension method. In practice there are several methods for calculating the SVD that are of higher accuracy and speed. Section 4 lists some references on the mathematics and computation of SVD.

IV. Implementation and Findings

Given two data values f1 and f2 from different QRRs, we require their similarity, s12, to be a real value in [0, 1]. The data value similarity is calculated according to the data type tree shown in Fig. 4. Each child node is a subset of its parent node. For example, the "string" type includes several children data types, which are common on the Web such as "datetime", "float" and "price". The maximum depth of the data type tree is 4. In the following, we will refer to a non-string data type as a specific data type. Given two data values f1 and f2, we first judge their data types and then fit them as deeply as possible into the nodes n1 and n2 of the data type tree. For example, given a string "784", we will put it in node "integer".

The similarity s12 between two data values f1 and f2 with data type nodes n1 and n2 is defined as:

$$s_{12} = \begin{cases} 0.5 & n_1 = p(n_2) \,\&\, n_1 \neq String \quad OR \quad n_2 = p(n_1) \,\&\, n_2 \neq String \\ 1 & n_1 = n_2 \neq String \\ cosine\ \ similarity & n_1 = n_2 = String \\ 0 & otherwise \end{cases}$$

where p(ni) refers to the parent node of ni in the data type tree. The similarity between data values f1 and f2 is set to:

- 0.5, if they belong to different specific data types that have a common parent.
- 1, if they belong to the same specific data type.

- string cosine similarity of f1 and f2, if both f1 and f2 belong to the string data type.
- 0 otherwise, which occurs when one of f1 and f2 belongs to the string data type and the other one belongs to a specific data type, or f1 and f2 belong to different specific data types without any direct parent.

V. Conclusion & Future Enhancements

a) Conclusion

Social tagging systems provide recommendations to users based on what tags other users have used on items. In this paper, we developed a unified framework to model the three types of entities that exist in a social tagging system: users, items, and tags. We examined multiway analysis on data modeled as 3-order tensor, to reveal the latent semantic associations between users, items, and tags. The multiway latent semantic analysis and dimensionality reduction is performed by combining the HOSVD method with the Kernel-SVD smoothing technique. Our approach improves recommendations by capturing users multimodal perception of item/tag/user. Moreover, we study a problem of how to provide user recommendations, which can have significant applications in real systems but which have not been studied in depth so far in related research. We also performed experimental comparison of the proposed method against state-of the-art recommendations algorithms, with two real data sets (Last.fm and BibSonomy). Our results show significant improvements in terms of effectiveness measured through recall/precision. As future work, we intend to examine different methods for extending SVD to high-order tensors such as the Parallel Factor Analysis. We also indent to apply different weighting methods for the initial construction of a tensor. A different weighting policy for the tensor's initial values could improve the overall performance of our approach.

b) Future Enhancements

Although SVD has been shown to be an accurate data extraction method, it still suffers from some limitations. First, it requires at least two QRRs in the query result page. Second, any optional attribute that appears as the start node in a data region will be treated as auxiliary information. Third, similar to other related works, SVD mainly depends on tag structures to discover data values. Therefore, Finally, as previously mentioned, if a query result page has more than one data region that contains result records and the records in the different data regions are not similar to each other, then SVD will select only one of the data regions and discard the others.

References Références Referencias

1. E. Acar and B. Yener, "Unsupervised Multiway Data Analysis: A Literature Survey," IEEE Trans. Knowledge and Data Eng., vol. 21, no. 1, pp. 6-20, Jan. 2009.
2. N. Ali-Hasan and A. Adamic, "Expressing Social Relationships on the Blog through Links and Comments," Proc. Int'l Conf. Weblogs and Social Media (ICWSM), 2007.
3. M. Berry, S. Dumais, and G. O'Brien, "Using Linear Algebra for Intelligent Information Retrieval," SIAM Rev., vol. 37, no. 4, pp. 573-595, 1994.
4. M. Brand, "Incremental Singular Value Decomposition of Uncertain Data with Missing Values," Proc. European Conf. Computer Vision (ECCV '02), 2002.
5. J. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proc. Conf. Uncertainty in Artificial Intelligence, pp. 43-52, 1998.
6. D. Buttler, L. Liu, and C. Pu, "A Fully Automated Object Extraction System for the World Wide Web," Proc. 21st Int'l Conf. Distributed Computing Systems, pp. 361-370, 2001.
7. K. C.-C. Chang, B. He, C. Li, M. Patel, and Z. Zhang, "Structured Databases on the Web: Observations and Implications," SIGMOD Record, vol. 33, no. 3, pp. 61-70, 2004.
8. C.H. Chang and S.C. Lui, "IEPAD: Information Extraction Based on Pattern Discovery," Proc. 10th World Wide Web Conf., pp. 681-688, 2001.
9. L. Chen, H.M. Jamil, and N. Wang, "Automatic Composite Wrapper Generation for Semi-Structured Biological Data Based on Table Structure Identification," SIGMOD Record, vol. 33, no. 2, pp. 58-64, 2004.
10. W. Cohen, M. Hurst, and L. Jensen, "A Flexible Learning System for Wrapping Tables and Lists in HTML Documents," Proc. 11th World Wide Web Conf., pp. 232-241, 2002.
11. W. Cohen and L. Jensen, "A Structured Wrapper Induction System for Extracting Information from Semi-Structured Documents," Proc. IJCAI 2001 Workshop Adaptive Text Extraction and Mining, 2001.
12. V. Crescenzi, G. Mecca, and P. Merialdo, "Roadrunner: Towards Automatic Data Extraction from Large Web Sites," Proc. 27th Int'l Conf. Very Large Data Bases, pp. 109-118, 2001.
13. D.W. Embley, D.M. Campbell, Y.S. Jiang, S.W. Liddle, D.W. Lonsdale, Y.-K. Ng, and R.D. Smith, "Conceptual-model-based Data Extraction from Multiple-record Web Pages," Data and Knowledge Engineering, vol. 31, no. 3, pp. 227-251, 1999.
14. A. V. Goldberg, R. E. Tarjan. "A new approach to the maximum flow problem", Proceedings of the

eighteenth annual ACM symposium on Theory of computing, pp. 136–146, 1986.

15. D. Gusfield, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology. Cambridge University Press, 1997.
16. B. Liu, W. S. Lee, P. S. Yu, and X. Li *proposed a Faster retrieval method for* Improved search engines are substantially faster than old ones. In some cases they have been demonstrated to be 1000 times faster.

User-centred design : Choosing between formats was not the only step the user had to take in older search engines. In addition, users had to choose between file name search or full text search, and also optionally specify the time the file was recently modified. To achieve a reasonable retrieval time, the user needed to input more information in order for the computer to do less, a feature which reflects a machine-oriented design. Newer search engines' retrieval speed allows them to reduce the query launching steps and complications to a minimum.

Incremental Search : One advantage of newer search engines is that they support incremental search, so that the search begins as soon as the user types the first character of the query. This has the benefit of being interactive: allowing users to refine their query in light of the results returned, and truncate the query after typing just a few characters if the target item is already in view. Older search engines were less efficient: prompting the user via form filling to specify multiple attribute fields and hit carriage return before the query is sent off. Incrementality, according to Raskin, has several advantages: (a) user and computer do not have to wait for each other, (b) users know they have typed enough to disambiguate their query because the desired file appears in the display, (c) users receive constant feedback as to the results of the search – they can correct spelling mistakes or refine search words without interrupting the search.

17. W. Ng, L. Deng, and D. L. Lee, proved that given these improvements in desktop search engines, it is now time to examine their implications: What are users' file retrieval preferences, what motivates retrieval by search, and what is the effect of improved desktop search engines on file retrieval preferences and file organization?

If the availability of these improved desktop search engines leads to a substantial increase in search, then it is reasonable to assume that this effect will continue to grow as search engines improve. If, on the other hand, no such effect is found, it raises questions regarding claims that improved search engines affect retrieval preferences and file organization, though it always can be claimed that future improvements in search could change this. As search engines are consistently improving and will continue to do so, the examination of their implications on PIM should be a continuous effort.

This page is intentionally left blank

Global Journal of Computer Science and Technology
Network, Web & Security
Volume 13 Issue 12 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Path-Constrained Data Gathering Scheme

By Khaled Almiani, Ahmad A. Twaissi, Mohammed A. Abuhelaleh, Bassam A. Alqaralleh & Albara Awajan

Al-Hussein Bin Talal University, Jordan

Abstract - Several studies in recent years have considered the use of mobile elements for data gathering in wireless sensor networks, so as to reduce the need for multi-hop forwarding among the sensor nodes and thereby prolong the network lifetime. Since, typically, practical constraints preclude a mobile element from visiting all nodes in the sensor network, the solution must involve a combination of a mobile element visiting a subset of the nodes (cache points), while other nodes communicate their data to the cache points wirelessly. This leads to the optimization problem of minimizing the communication distance of the sensor nodes, while keeping the tour length of the mobile element below a given constraint. In this paper, we investigate the problem of designing the mobile elements tours such that the length of each tour is below a per-determined length and the number of hops between the tours and the nodes not included in the tour is minimized. To address this problem, we present an algorithmic solution that consider the distribution of the nodes during the process of building the tours. We compare the resulting performance of our algorithm with the best known comparable schemes in the literature.

GJCST-E Classification : G.2.2

Strictly as per the compliance and regulations of:

Path-Constrained Data Gathering Scheme

Khaled Almiani [α], Ahmad A. Twaissi [σ], Mohammed A. Abuhelaleh [ρ], Bassam A. Alqaralleh [ω] & Albara Awajan[¥]

Abstract - Several studies in recent years have considered the use of mobile elements for data gathering in wireless sensor networks, so as to reduce the need for multi-hop forwarding among the sensor nodes and thereby prolong the network lifetime. Since, typically, practical constraints preclude a mobile element from visiting all nodes in the sensor network, the solution must involve a combination of a mobile element visiting a subset of the nodes (cache points), while other nodes communicate their data to the cache points wirelessly. This leads to the optimization problem of minimizing the communication distance of the sensor nodes, while keeping the tour length of the mobile element below a given constraint. In this paper, we investigate the problem of designing the mobile elements tours such that the length of each tour is below a per-determined length and the routing trees size is minimized. To address this problem, we present an algorithmic solution that consider the distribution of the nodes during the process of building the tours. We compare the resulting performance of our algorithm with the best known comparable schemes in the literature.

I. Introduction

Many typical applications of wireless sensor networks (WSNs) involve the collection of data obtained by sensor nodes at a pre-defined sink. This is normally achieved by wireless transmission of the data, possibly over several hops (especially in applications where the sensors are deployed in a hostile or hard-to-access environment). In many cases, the wireless communication results in a major energy expenditure that limits the operational lifetime of the network. Even worse, in multi-hop scenarios, the depletion of the sensors' energy sources (such as batteries) is non-uniform, as nodes that are close to the sink are required to forward all the data traffic and are likely to be the first to run out of energy. Once these sensors fail, other nodes can no longer reach the sink, and the network ceases to operate even though ample energy remains at nodes further away from the sink. This common problem occurs largely independently of the communication protocols used in the network.

In general, the use of Mobile Elements (MEs) [1], [2], [3] can significantly increases the lifetime of the network. Mobile element roams in the network and collects data from sensors via short range communication, the energy cost of which is considerably lower. Thus, the lifetime of the network increases by avoiding multi-hop communication. The main drawback for this approach is the increased latency of the data collection. Typically, the speed of mobile element can be about 0:1-2 m/s [4], [5], resulting in substantial traveling time for the ME and, correspondingly, delay in gathering the sensors' data. In practice, often the ME tour length is bounded by a predetermined time deadline, either due to timeliness constraints on the sensor data or a limit on the amount of energy available to the ME itself. A possible solution is to employ more than one ME; however, this solution is often impractical due to the high cost of MEs, and may not in fact help at all if some sensors are beyond reach due to ME battery limitations in the first place.

To address this problem, several proposals presented a hybrid approach, which combines multi-hop forwarding with the use of mobile elements. In this approach, mobile element visits subset of the nodes termed as caching points. These caching points stores the data of the nodes that are not included in the tour of the mobile element. Once a mobile element become within the transmission range of a caching point, the caching point transmit its data to the mobile element. By adopting such an approach, the mobile element gather the data of the entire without the need of visiting each node physically. In this direction, we investigate the problem of designing the tours for the mobile elements and the data forwarding trees, with the objective of minimizing the distance between nodes not included in the tour and the tour itself. We propose a heuristic-based solution that creates its solution by partition the network into clusters. The in each cluster a tour will be constructed to satisfy the objective. The results show that our scheme significantly outperforms the best comparable scheme in the literature.

The rest of the paper is organized as follows. Section 2 presents the related work in this research area. Section 3 presents the Problem definition. In Section 4, we present the details of our algorithmic solution. Section 5 presents the evaluation. Finally, Section 6 concludes the paper.

II. Related Work

There have been many proposals in recent literature that studied using mobile element(s) to prolong the lifetime of the network. Based on the categorization given in [3], we review three major approaches.

Authors α σ ρ ω : Al-Hussein Bin Talal University, Jordan.
E-mails: twaissi@ahu.edu.jo, mabuhela@ahu.edu.jo, alqaralleh@ahu.edu.jo, k.almiani@ahu.edu.jo.
Author ¥: Albalqa Applied University, Jordan.
E-mail : a.awajan@bau.edu.jo

In a typical flat-topology network, the nodes around the sink are likely to be the first to die due to having to forward the data traffic from all other sensors. Based on this observation, several proposals [6], [7] have investigated using mobile sink(s) to reduce the energy consumption in the network. By varying the path to the sink(s), the residual energy in the nodes becomes more evenly balanced throughout the network, leading to a higher network's lifetime. However, in order to be effective, this technique requires the sink location (and routes thereto) to change regularly, which places a potentially prohibitive overhead on the nodes due to the frequent re-computation of the routes. Zhao et al [8], [9] investigated the problem of maximizing the overall network utility. Accordingly, they presented two distributed algorithms for data gathering where the mobile sink stays at each anchor point (gathering point) for a period of sojourn time and collects data from nearby sensors via multi-hop communications. They considered the cases where the sojourn time is fixed as well as variable.

In the second approach, mobile elements travel across the network and gather each sensors data via single-hop, short-range communication. In this scenario, the problem of computing the ME tours is exactly the Traveling Salesman Problem (TSP) [10], with the possibility of adding additional constraints to capture the limitations of the nodes buffer size. In [11], [12], [13], [14], [15] several heuristics have been proposed to that effect, so that each sensor is visited before its buffer is full. Although this approach substantially reduces the energy consumption by avoiding multi-hop communications, it incurs a high delay when the network area is large, because of the requirement that the MEs physically visit all sensor nodes.

Finally, the third approach is a hybrid that combines multihop forwarding with data collection by mobile elements. Our work falls into this category. Some earlier works, e.g. [16], [17], [18], assumed the mobile route to be predetermined and were mainly concerned with the timing of transmissions, aiming to minimize the need for in-network caching by timing the transmissions to coincide with the passing of the tour. In [19], [20], the minimum-energy Rendezvous Planning Problem (RPP) is introduced. This problem deals with determining the set of rendezvous points constructing the ME tour. In RPP, the goal is to minimize the Euclidean distance between the source nodes and the tour. Path finding algorithms based on maxflow computations have been considered by [21]. In that work, the authors use a standard maxflow formulation to represent the sensor network. However the problem they consider is finding a path through anywhere in the network area, which does not need to move from a sensor location to another sensor location.

Xu et al [22] also proposed a tour finding algorithm in which nodes away from the caching points send their data to the caching points using multi-hop communication. The main concept of their algorithm is to find a tour that satisfies the transit constraint such that the depth of the routing trees connected to this tour are bounded by pre-determined variable h. this algorithm starts by setting the value of h to 1 and increasing it gradually, until such a tour is found. By bounding the depth of the routing tress, the algorithm aims to reduce the energy consumption due to multi-hop forwarding. However, important factors in determining the lifetime of the network such as structure of routing tress, the energy level of the nodes, and the distribution of the caching points were not mainly considered in this algorithm. Liang et al [23] investigated similar problem where the depth of the routing trees were bounded by pre-determined variable. The problem presented in this work shares some similarities with the Vehicle Routing Problem (VRP) [24]. Given a fleet of vehicles assigned to a depot, VRP deals with determining the fleet routes to deliver goods from a depot to customers while minimizing the vehicles' total travel cost.

III. Problem Definition

We are given an undirected graph $G = (V, E)$, where V is the set of vertices representing the locations of the sensors in the network, and E is the set of edges that represents the communication network topology, i.e. $(v_i, v_j) \in E$ if and only if v_i and v_j are within each others communication range. In addition, we are given k, that represents the number of tours need to be constructed. The complete graph $G' = (V, E')$, where $E' = V \times V$, represents the possible movements of the mobile elements. Each edge $(v_i, v_j) \in E'$ has a length r_{ij}, which represents the time needed by a mobile element to travel between sensor v_i and v_j. The data of all sensors must be uploaded to a mobile element periodically at least once in L time units, where L is determined from the application requirements and the sensors buffer size. In other words, we assume that each mobile element conducts its tour periodically, with L being a constraint on the maximum tour length. In this paper, for simplicity, we assume that the mobile element travels at constant speed, and that, therefore, the travelling times between sensors (r_{ij}) correspond directly to their respective Euclidean distances; however, this assumption is not essential to our algorithms and can be easily dropped if necessary.

In our problem, we seek to find the k tours, where the length of the tour is bounded by L, such that the number of hops between any node and its caching point is minimized.

IV. Tours and Forwarding Trees

The goal of our algorithm is to find the tours for the given k mobile elements such that distance between the nodes not included in the tours and the tours are

minimized. In this direction, we first partition the network into k partitions. The underling goal for the partitioning processes is to minimize the distance between nodes belong to the same cluster. By adopting such a process, we aim to minimize the distance between the nodes and the tours in each cluster. Once the clusters are constructed, the tour building step works to create a tour in each cluster with the aim of minimizing the total distance of the routing trees.

a) Clustering Step

The clustering step attempts to find a given number of clusters such that the sum of hop-distances among nodes belonging to the same cluster is minimized. The clustering process start by selecting a node randomly as a cluster centroid. Once a centroid node is identified it will be added to list termed R. Then, the process works by identifying k-1 cluster centroids, where in each iteration, a node is identify as centroid if it has the maximum total hop-distance to all nodes stored in R. Once all k clusters are identified, each node not chosen

	Input: G (topology graph), c (number of clusters to be established)
	Output: a set of clusters
1	Randomly choose a centroid centers and add it to R
2	**Do**
3	for all nodes in G
4	calculate the distance to all nodes in R (in terms of hop-distance)
5	Select the node with the maximum distance as new centroid
6	Add the select node to R
7	Until k centroid is selected
8	assign each node to its nearest centroid
9	identify each centroid and the nodes assigned to it as a cluster

Figure 1 : The clustering step

a centroid will be assign to its nearest cluster centroid. By adopting such a mechanism, we aim to direct the partitioning to group the nodes into clusters based on their distribution. Figure 1 outlines the process of this step.

b) Caching point identification step

Now, in each cluster subset of nodes will be selected as caching points. These caching points will store the other nodes data and will be used to construct the mobile element tour. In each cluster, this process works by first identifying the center node. The center node is defined as the node that has the minimum total hop-distance to all other nodes in the cluster. Then, it works to constructs Minimum Spanning Tree (MST) rooted at the cluster center node. Once this MST is constructed the process proceeds to traverse the constructed tree using BST mechanism. This traversing stops once the total distance of the visited edges reach (L. 2/3). As we will see in the tours constructing step, this condition depends on the employed TSP algorithm. The last step in this process is to identify the visited nodes in the traversing mechanism as caching points.

c) Constructing the tour

The tour construction phase uses the nodes identified as caching points in the previous step and can be based on any TSP algorithm or heuristic. We use the Christofides approximation algorithm here, as it is known algorithm with 2/3 approximation factor.

V. Experimental Evaluation

To evaluate the presented algorithm's performance, we conducted an extensive set of experiments using the J-sim simulator [25]. We used the following parameters:

(1) The area of the network is 250,000m^2. (2) The tour length constraint L is set to $0{:}05 \cdot s \cdot T_L$, where $s = 1$ m/s is the speed of the mobile element, and T_L is the total length of the edges in the minimum spanning tree that connects all nodes, for 500 nodes in the network. (3) The starting value of M is chosen to be $0{:}5 \cdot T_H$, where T_H is the number of hops between the farthest nodes in the network. (4) The radio parameters are set according to the MICAz data sheet [26], namely: the radio bandwidth is 250 Kbps, the transmission power is 21 mW, the receiving power is 15 mW, and the initial battery power is 10 Joules. For simplicity, we only account for the radio receiving and transmitting energy.

We are particularly interested in investigating the performance of the presented algorithm in terms of the lifetime of the network and the total distance of the routing trees.

We consider the following deployment scenarios:

1. Uniform density deployment: in this scenario, we assume that the nodes are uniformly deployed in a square area of 500×500m^2.
2. Variable density deployment: in this scenario, we divide the network into a 10×10 grid of squares,

where each square is $50\times50m^2$. We randomly choose 30 of the squares, and in each one of those we fix the node density to be x times the density in the remaining squares. x is a density parameter, which in most experiments (unless mentioned otherwise) is set to $x = 5$.

We compare our algorithm to a modified version of the FFT algorithm [27], we refer to this version as V-FFT. In the original FFT algorithm, a new node is added to the tour based on benefit function. The benefit value of each node depends on the distance between this node and the currently constructed tour as well as the number of nodes it covers. The number of nodes that are covered by any tour node is controlled by the parameter $h \geq 1$. This parameter refers to the maximum number of hops allowed between any two nodes: a tour node can cover any node at most h hops away. Initially, $h = 1$ and in each round the value of h is incremented by one until a tour that satisfies the transit constraint is determined. In the modified version of the FFT algorithm (V-FFT), we fix h to be 0:5 the maximum distance between any two nodes inside any cluster obtained by our heuristic. Then, we start by constructing the first tour. Each selected caching point and the neighbor of this caching points, will be removed from consideration at later stage. This tour will be extended, based on the given cost function, until the constructing tour cannot be extended without violating the transit constraint. Once such a tour is obtained, a new tour will be constructed using the same mechanism.

We evaluate the impact of the number of nodes on the lifetime of the network and the total size of the routing trees each algorithm obtains. Figure 2, Figure 3, Figure 5 and Figure 5 show the results for deployment scenarios. In the

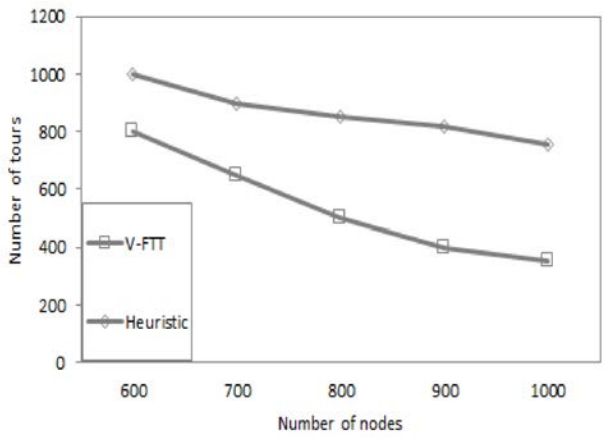

Figure 2 : Network lifetime against the number of nodes, for the uniform density deployment scenario

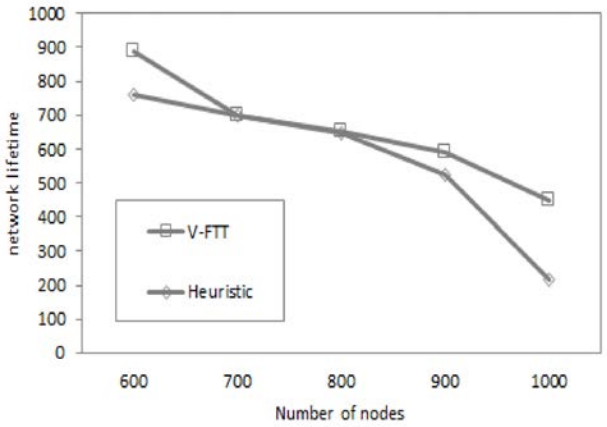

Figure 3 : Network lifetime against the number of nodes, for the variable density deployment scenario

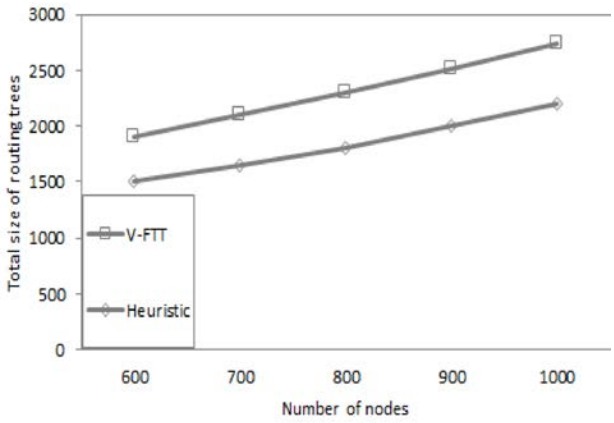

Figure 4 : Total size of routing trees against the number of nodes, for the uniform density deployment scenario

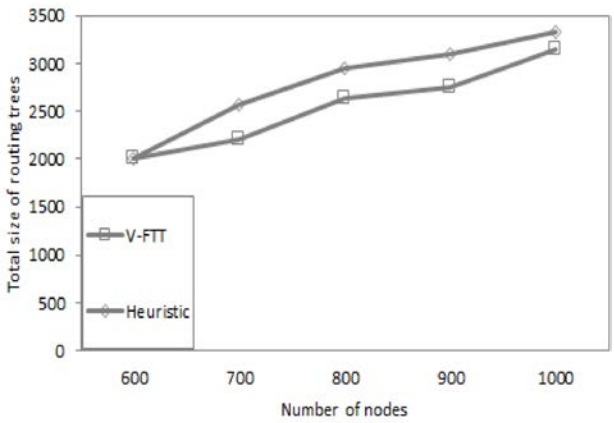

Figure 5 : Total size of routing trees against the number of nodes, for the variable density deployment scenario

uniform deployment scenario, we can see that increasing the number of nodes increases the gap between the algorithm performances. In this deployment

scenario, since we deal with uniformly deployment scenario, the locality of each cluster obtained by the proposed algorithm is expected to be the same. Such behavior results in creating a relatively linear relationship between network lifetime and the number of node. This can be noticed in the results of the routing trees size experiments. In the V-FFT algorithm, the benefit function takes into account the number of nodes covered by the considered nodes as well as the distance between this node and the current constructed tour. And considering the number of covered-nodes as well as the stochastic behavior of such benefit function are the factor behind the seen performance.

In the variable deployment scenario, we can see that increasing the number of nodes results in slightly reducing the gap between the algorithms performances. To understand this behavior, let us discuss the main mechanism behind each algorithm performance. In the proposed algorithm, the obtained clusters is expected to be centralized at the dense grids. This is expected to significantly reduce the lifetime of the network, since in each cluster the tour will be saturated with nodes very closed to each other. This become more obvious while increasing the number of tours. In the V-FFT algorithm, as we mentioned, the benefit function take into consideration the number of neighboring nodes covered by a node as well as the distance between the node and the current constructed tour. In this deployment scenario, considering the number of neighboring nodes during the construction of the tours is expected to improve the V-FFT performance, since it will avoid adding caching node that is very closed to the current constructed tour. These can mainly describe the seen performance.

VI. Conclusions

In this paper, we consider the problem of designing the mobile elements tours such that total size of the routing trees is minimized. In this work, we present an algorithmic solution that create its solution by partitioning the network, then in each clusters, a tour will be constructed based on the distribution of the nodes.

An interesting open problem would be to consider application scenarios where the data gathering latency requirements vary in the network. For example, some areas in the network need to send data more frequently than others. In this case the tour length constraints would be different for different areas.

References Références Referencias

1. J. Butler, "Robotics and Microelectronics: Mobile Robots as Gateways into Wireless Sensor Networks," *Technology@Intel Magazine,* 2003.
2. A. LaMarca, W. Brunette, D. Koizumi, M. Lease, S. Sigurdsson, K. Sikorski, D. Fox, and G. Borriello, "Making sensor networks practical with robots," *Lecture notes in computer science,* pp. 152–166, 2002. [Online]. Available: http://www.springerlink.com/index/9BUYG5K8BEW05CRX.pdf
3. E. Ekici, Y. Gu, and D. Bozdag, "Mobility-based communication in wireless sensor networks," *IEEE Communications Magazine,* vol. 44, no. 7, pp. 56–62, 2006.
4. K. Dantu, M. Rahimi, H. Shah, S. Babel, A. Dhariwal, and G. Sukhatme, "Robomote: enabling mobility in sensor networks," in *Proceedings of the 4th international symposium on Information processing in sensor networks (IPSN).* IEEE Press, 2005, pp. 404–409. [Online]. Available: http://portal.acm.org/citation.cfm?id=1147685.1147751
5. R. Pon, M. Batalin, J. Gordon, A. Kansal, D. Liu, M. Rahimi, L. Shirachi, Y. Yu, M. Hansen, W. Kaiser, and Others, "Networked infomechanical systems: a mobile embedded networked sensor platform," in *Proceedings of the 4th international symposium on Information processing in sensor networks.* IEEE Press, 2005, pp. 376–381. [Online]. Available: http://portal.acm.org/citation.cfm?id=1147685.1147746
6. S. Gandham, M. Dawande, R. Prakash, and S. Venkatesan, "Energy efficient schemes for wireless sensor networks with multiple mobile base stations," in *Proceedings of IEEE Globecom,* vol. 1. IEEE, 2003, pp. 377–381.
7. Z. Wang, S. Basagni, E. Melachrinoudis, and C. Petrioli, "Exploiting sink mobility for maximizing sensor networks lifetime," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS),* vol. 9, 2005, pp. 03–06. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/HICSS.2005.259
8. M. Zhao and Y. Yang, "Optimization-Based Distributed Algorithms for Mobile Data Gathering in Wireless Sensor Networks," *IEEE Transactions on Mobile Computing,* vol. 11, no. 10, pp. 1464–1477, 2012.
9. ——, "Efficient data gathering with mobile collectors and space-division multiple access technique in wireless sensor networks," *IEEE Transactions on Computers,* vol. 60, no. 3, pp. 400–417, 2011.
10. N. Christofides, "Worst-case analysis of a new heuristic for the traveling salesman problem," 1976.
11. A. Somasundara, A. Ramamoorthy, and M. Srivastava, "Mobile element scheduling with dynamic deadlines," *IEEE transactions on Mobile Computing,* vol. 6, no. 4, pp. 395–410, 2007. [Online]. Available: http://www.ece.iastate.edu/~adityar/adi docs/04116703.pdf
12. A. Somasundara, A. Ramamoorthy, and B. Srivastava, "Mobile Element Scheduling for Efficient Data Collection in Wireless Sensor Networks with Dynamic Deadlines," in *Real-Time Systems Symposium, 2004. Proceedings. 25th IEEE International.* IEEE Press, 2004, pp. 296–305.

13. Y. Gu, D. Bozdag, E. Ekici, F. Ozguner, and C. Lee, "Partitioning based mobile element scheduling in wireless sensor networks," in *Sensor and Ad Hoc Communications and Networks, 2005. IEEE SECON 2005. 2005 Second Annual IEEE Communications Society Conference on,* 2005, pp. 386–395.
14. K. Almi'ani, S. Selvadurai, and A. Viglas, "Periodic Mobile Multi-Gateway Scheduling," in *Proceedings of the Ninth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT).* IEEE, 2008, pp. 195–202. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4710981
15. M. Ma and Y. Yang, "Data gathering in wireless sensor networks with mobile collectors," in *2008 IEEE International Symposium on Parallel and Distributed Processing.* IEEE, Apr. 2008, pp. 1–9. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4536269
16. A. Somasundara, A. Kansal, D. Jea, D. Estrin, and M. Srivastava, "Controllably mobile infrastructure for low energy embedded networks," *IEEE Transactions on Mobile Computing,* vol. 5, no. 8, pp. 958–973, 2006.
17. R. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling a Three-tier Architecture for Sparse Sensor Networks," in *Proceedings of the First IEEE Workshop on Sensor Network Protocols and Applications,* 2003, pp. 30–41. [Online]. Available: http://www.comp.nus.edu.sg/~tayyc/6282/forPresentation/SensorNet.pdf
18. D. Jea, A. Somasundara, and M. Srivastava, "Multiple controlled mobile elements (data mules) for data collection in sensor networks," *Lecture Notes in Computer Science,* vol. 3560, pp. 244–257, 2005. [Online]. Available: http://www.springerlink.com/index/4617cluarmakkx5x.pdf
19. G. Xing, T. Wang, Z. Xie, and W. Jia, "Rendezvous planning in wireless sensor networks with mobile elements," *IEEE Transactions on Mobile Computing,* vol. 7, no. 12, pp. 1430–1443, Dec. 2008. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4492781
20. G. Xing, T. Wang, W. Jia, and M. Li, "Rendezvous design algorithms for wireless sensor networks with a mobile base station," in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc).* ACM New York, NY, USA, 2008, pp. 231–240. [Online]. Available: http://portal.acm.org/citation.cfm?id=1374650
21. M. Ma and Y. Yang, "SenCar: An Energy-Efficient Data Gathering Mechanism for Large-Scale Multihop Sensor Networks," *IEEE Transactions on Parallel and Distributed Systems,* vol. 18, no. 10, pp. 1476–1488, Oct. 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4302733
22. Z. Xu, W. Liang, and Y. Xu, "Network Lifetime Maximization in Delay-Tolerant Sensor Networks With a Mobile Sink," in *Proceedings of the 8th IEEE International Conference on Distributed Computing in Sensor Systems,* 2012, pp. 9–16.
23. W. Liang and J. Luo, "Network Lifetime Maximization in Sensor Networks with Multiple Mobile Sinks," in *Proceedings of the 36th IEEE Conference on Local Computer Networks (LCN).* IEEE, 2011, pp. 350–357.
24. P. Toth and D. Vigo, "The Vehicle Routing Problem," *Society for Industrial & Applied Mathematics (SIAM),* 2001.
25. J. Hou, L. Kung, N. Li, H. Zhang, W. Chen, H. Tyan, and H. Lim, "J-Sim: A Simulation and emulation environment for wireless sensor networks," *IEEE Wireless Communications Magazine,* vol. 13, no. 4, pp. 104–119, 2006.
26. J. Hill and D. Culler, "Mica: A wireless platform for deeply embedded networks," *IEEE micro,* vol. 22, no. 6, pp. 12–24, 2002.
27. Z. Xu, W. Liang, and Y. Xu, "Network Lifetime Maximization in Delay-Tolerant Sensor Networks With a Mobile Sink," in *Proceedings of the 8th IEEE International Conference on Distributed Computing in Sensor Systems,* 2012, pp. 9–16.

Global Journals Inc. (US) Guidelines Handbook 2013

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

- 'FARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FARSC" can be added to name in the following manner. eg. **Dr. John E. Hall, Ph.D., FARSC or William Walldroff Ph. D., M.S., FARSC**
- Being FARSC is a respectful honor. It authenticates your research activities. After becoming FARSC, you can use 'FARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 60% Discount will be provided to FARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- FARSC will be given a renowned, secure, free professional email address with 100 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- FARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 15% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- Eg. If we had taken 420 USD from author, we can send 63 USD to your account.
- FARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- After you are FARSC. You can send us scanned copy of all of your documents. We will verify, grade and certify them within a month. It will be based on your academic records, quality of research papers published by you, and 50 more criteria. This is beneficial for your job interviews as recruiting organization need not just rely on you for authenticity and your unknown qualities, you would have authentic ranks of all of your documents. Our scale is unique worldwide.
- FARSC member can proceed to get benefits of free research podcasting in Global Research Radio with their research documents, slides and online movies.
- After your publication anywhere in the world, you can upload you research paper with your recorded voice or you can use our professional RJs to record your paper their voice. We can also stream your conference videos and display your slides online.
- FARSC will be eligible for free application of Standardization of their Researches by Open Scientific Standards. Standardization is next step and level after publishing in a journal. A team of research and professional will work with you to take your research to its next level, which is worldwide open standardization.

- FARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), FARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 80% of its earning by Global Journals Inc. (US) will be transferred to FARSC member's bank account after certain threshold balance. There is no time limit for collection. FARSC member can decide its price and we can help in decision.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

- 'MARSC' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'MARSC" can be added to name in the following manner. eg. Dr. John E. Hall, Ph.D., MARSC or William Walldroff Ph. D., M.S., MARSC
- Being MARSC is a respectful honor. It authenticates your research activities. After becoming MARSC, you can use 'MARSC' title as you use your degree in suffix of your name. This will definitely will enhance and add up your name. You can use it on your Career Counseling Materials/CV/Resume/Visiting Card/Name Plate etc.
- 40% Discount will be provided to MARSC members for publishing research papers in Global Journals Inc., if our Editorial Board and Peer Reviewers accept the paper. For the life time, if you are author/co-author of any paper bill sent to you will automatically be discounted one by 60%
- MARSC will be given a renowned, secure, free professional email address with 30 GB of space eg.johnhall@globaljournals.org. You will be facilitated with Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.
- MARSC member is eligible to become paid peer reviewer at Global Journals Inc. to earn up to 10% of realized author charges taken from author of respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account or to your PayPal account.
- MARSC member can apply for free approval, grading and certification of some of their Educational and Institutional Degrees from Global Journals Inc. (US) and Open Association of Research, Society U.S.A.
- MARSC is eligible to earn from their researches: While publishing his paper with Global Journals Inc. (US), MARSC can decide whether he/she would like to publish his/her research in closed manner. When readers will buy that individual research paper for reading, 40% of its earning by Global Journals Inc. (US) will be transferred to MARSC member's bank account after certain threshold balance. There is no time limit for collection. MARSC member can decide its price and we can help in decision.

AUXILIARY MEMBERSHIPS

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJCST for one year (subscription for one year).
- The member will be allotted free 1 GB Web-space along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted free 500 MB email space.

PAPER PUBLICATION

- The members can publish paper once. The paper will be sent to two-peer reviewer. The paper will be published after the acceptance of peer reviewers and Editorial Board.

Process of submission of Research Paper

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

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

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

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

(II) Choose corresponding Journal.

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

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

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

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

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

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

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

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

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

Scope

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

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

2. ETHICAL GUIDELINES

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

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

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

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

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

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

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

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

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

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

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

Please mention proper reference and appropriate acknowledgements wherever expected.

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

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

3. SUBMISSION OF MANUSCRIPTS

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

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

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

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

4. MANUSCRIPT'S CATEGORY

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

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

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

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

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

5. STRUCTURE AND FORMAT OF MANUSCRIPT

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

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

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

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

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

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

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

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

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

(h) Brief Acknowledgements.

(i) References in the proper form.

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

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

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

Format

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

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

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

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

Structure

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

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

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

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

Key Words

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

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

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

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

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

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

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

Acknowledgements: Please make these as concise as possible.

References

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

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

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

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

Tables, Figures and Figure Legends

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

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

Preparation of Electronic Figures for Publication

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

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

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

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

6. AFTER ACCEPTANCE

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

6.1 Proof Corrections

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

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

(Free of charge) from the following website:

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

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

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

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

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

6.3 Author Services

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

6.4 Author Material Archive Policy

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

6.5 Offprint and Extra Copies

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

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

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

TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Informal Guidelines of Research Paper Writing

Key points to remember:

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

Final Points:

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

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

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

General style:

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

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

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

In every sections of your document

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

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

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

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

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

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

Title Page:

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

Abstract:

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

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

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

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

Approach:

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

Introduction:

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

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

Approach:

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

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

Procedures (Methods and Materials):

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

Materials:

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

Methods:

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

Approach:

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

What to keep away from

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

Results:

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

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

Content

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

What to stay away from

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

Approach

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

Figures and tables

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

Discussion:

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

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

Approach:

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

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

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

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

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

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

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

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

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

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

INDEX

save our planet

Global Journal of Computer Science and Technology

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

ISSN 9754350