



IJCSI

International Journal of Computer Science Issues

Volume 10, Issue 4, No 2, July 2013
ISSN (Online): 1694-0784
ISSN (Print): 1694-0814

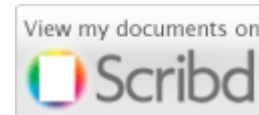
© IJCSI PUBLICATION
www.IJCSI.org

IJCSI proceedings are currently indexed by:



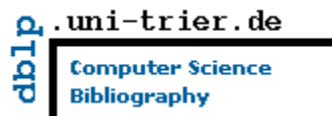
Cogprints

Google scholar



SciRate.com

CiteSeer^x beta



DOAJ DIRECTORY OF OPEN ACCESS JOURNALS



ProQuest

IJCSI Publicity Board 2013

Dr. Borislav D Dimitrov

Department of General Practice, Royal College of Surgeons in Ireland
Dublin, Ireland

Dr. Vishal Goyal

Department of Computer Science, Punjabi University
Patiala, India

Mr. Nehinbe Joshua

University of Essex
Colchester, Essex, UK

Mr. Vassilis Papataxiarhis

Department of Informatics and Telecommunications
National and Kapodistrian University of Athens, Athens, Greece

IJCSI Editorial Board 2013

Dr Tristan Vanrullen

Chief Editor

LPL, Laboratoire Parole et Langage - CNRS - Aix en Provence, France

LABRI, Laboratoire Bordelais de Recherche en Informatique - INRIA - Bordeaux, France

LEEE, Laboratoire d'Esthétique et Expérimentations de l'Espace - Université d'Auvergne, France

Dr Constantino Malagôn

Associate Professor

Nebrija University

Spain

Dr Lamia Fourati Chaari

Associate Professor

Multimedia and Informatics Higher Institute in SFAX

Tunisia

Dr Mokhtar Beldjehem

Professor

Sainte-Anne University

Halifax, NS, Canada

Dr Pascal Chatonnay

Assistant Professor

Maître de Conférences

Laboratoire d'Informatique de l'Université de Franche-Comté

Université de Franche-Comté

France

Dr Karim Mohammed Rezaul

Centre for Applied Internet Research (CAIR)

Glyndwr University

Wrexham, United Kingdom

Dr Yee-Ming Chen

Professor

Department of Industrial Engineering and Management

Yuan Ze University

Taiwan

Dr Gitesh K. Raikundalia

School of Engineering and Science,

Victoria University

Melbourne, Australia

Dr Vishal Goyal

Assistant Professor
Department of Computer Science
Punjabi University
Patiala, India

Dr Dalbir Singh

Faculty of Information Science And Technology
National University of Malaysia
Malaysia

Dr Natarajan Meghanathan

Assistant Professor
REU Program Director
Department of Computer Science
Jackson State University
Jackson, USA

Dr. Prabhat K. Mahanti

Professor
Computer Science Department,
University of New Brunswick
Saint John, N.B., E2L 4L5, Canada

Dr Navneet Agrawal

Assistant Professor
Department of ECE,
College of Technology & Engineering,
MPUAT, Udaipur 313001 Rajasthan, India

Dr Panagiotis Michailidis

Division of Computer Science and Mathematics,
University of Western Macedonia,
53100 Florina, Greece

Dr T. V. Prasad

Professor
Department of Computer Science and Engineering,
Lingaya's University
Faridabad, Haryana, India

Dr Saqib Rasool Chaudhry

Wireless Networks and Communication Centre
261 Michael Sterling Building
Brunel University West London, UK, UB8 3PH

Dr Shishir Kumar

Department of Computer Science and Engineering,
Jaypee University of Engineering & Technology
Raghogarh, MP, India

Dr P. K. Suri

Professor
Department of Computer Science & Applications,
Kurukshetra University,
Kurukshetra, India

Dr Paramjeet Singh

Associate Professor
GZS College of Engineering & Technology,
India

Dr Majid Bakhtiari

Faculty of Computer Science & Information System
University technology Malaysia
Skudai, 81310 Johore, Malaysia

Dr Shaveta Rani

Associate Professor
GZS College of Engineering & Technology,
India

Dr. Seema Verma

Associate Professor,
Department Of Electronics,
Banasthali University,
Rajasthan - 304022, India

Dr G. Ganesan

Professor
Department of Mathematics,
Adikavi Nannaya University,
Rajahmundry, A.P, India

Dr A. V. Senthil Kumar

Department of MCA,
Hindusthan College of Arts and Science,
Coimbatore, Tamilnadu, India

Dr Mashiur Rahman

Department of Life and Coordination-Complex Molecular Science,
Institute For Molecular Science, National Institute of Natural Sciences,
Miyodaiji, Okazaki, Japan

Dr Jyoteesh Malhotra

ECE Department,
Guru Nanak Dev University,
Jalandhar, Punjab, India

Dr R. Ponnusamy

Professor

Department of Computer Science & Engineering,
Aarupadai Veedu Institute of Technology,
Vinayaga Missions University, Chennai, Tamilnadu, India

Dr Nittaya Kerdprasop

Associate Professor

School of Computer Engineering,
Suranaree University of Technology, Thailand

Dr Manish Kumar Jindal

Department of Computer Science and Applications,
Panjab University Regional Centre, Muktsar, Punjab, India

Dr Deepak Garg

Computer Science and Engineering Department,
Thapar University, India

Dr P. V. S. Srinivas

Professor

Department of Computer Science and Engineering,
Geethanjali College of Engineering and Technology
Hyderabad, Andhra Pradesh, India

Dr Sara Moein

CMSSP Lab, Block A, 2nd Floor, Faculty of Engineering,
MultiMedia University, Malaysia

Dr Rajender Singh Chhillar

Professor

Department of Computer Science & Applications,
M. D. University, Haryana, India

EDITORIAL

In this forth edition of 2013, we bring forward issues from various dynamic computer science fields ranging from system performance, computer vision, artificial intelligence, software engineering, multimedia, pattern recognition, information retrieval, databases, security and networking among others.

Considering the growing interest of academics worldwide to publish in IJCSI, we invite universities and institutions to partner with us to further encourage open-access publications.

As always we thank all our reviewers for providing constructive comments on papers sent to them for review. This helps enormously in improving the quality of papers published in this issue.

Google Scholar reported a large amount of cited papers published in IJCSI. We will continue to encourage the readers, authors and reviewers and the computer science scientific community and interested authors to continue citing papers published by the journal.

Apart from availability of the full-texts from the journal website, all published papers are deposited in open-access repositories to make access easier and ensure continuous availability of its proceedings free of charge for all researchers.

We are pleased to present IJCSI Volume 10, Issue 4, No 2, July 2013 (IJCSI Vol. 10, Issue 4, No 2). The acceptance rate for this issue is 29.7%.



IJCSI Editorial Board
July 2013 Issue
ISSN (online): 1694-0784
ISSN (print): 1694-0814
© IJCSI Publications
www.IJCSI.org

IJCSI Reviewers Committee 2013

Mr. Markus Schatten, University of Zagreb, Faculty of Organization and Informatics, Croatia

Mr. Vassilis Papataxiarhis, Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens, Greece

Dr Modestos Stavrakis, University of the Aegean, Greece

Dr Fadi KHALIL, LAAS -- CNRS Laboratory, France

Dr Dimitar Trajanov, Faculty of Electrical Engineering and Information technologies, ss. Cyril and Methodius Univesity - Skopje, Macedonia

Dr Jinping Yuan, College of Information System and Management,National Univ. of Defense Tech., China

Dr Alexis Lazanas, Ministry of Education, Greece

Dr Stavroula Mougiakakou, University of Bern, ARTORG Center for Biomedical Engineering Research, Switzerland

Dr Cyril de Runz, CReSTIC-SIC, IUT de Reims, University of Reims, France

Mr. Pramodkumar P. Gupta, Dept of Bioinformatics, Dr D Y Patil University, India

Dr Alireza Fereidunian, School of ECE, University of Tehran, Iran

Mr. Fred Viezens, Otto-Von-Guericke-University Magdeburg, Germany

Dr. Richard G. Bush, Lawrence Technological University, United States

Dr. Ola Osunkoya, Information Security Architect, USA

Mr. Kotsokostas N.Antonios, TEI Piraeus, Hellas

Prof Steven Totosy de Zepetnek, U of Halle-Wittenberg & Purdue U & National Sun Yat-sen U, Germany, USA, Taiwan

Mr. M Arif Siddiqui, Najran University, Saudi Arabia

Ms. Ilknur Icke, The Graduate Center, City University of New York, USA

Prof Miroslav Baca, Faculty of Organization and Informatics, University of Zagreb, Croatia

Dr. Elvia Ruiz Beltrán, Instituto Tecnológico de Aguascalientes, Mexico

Mr. Moustafa Banbouk, Engineer du Telecom, UAE

Mr. Kevin P. Monaghan, Wayne State University, Detroit, Michigan, USA

Ms. Moira Stephens, University of Sydney, Australia

Ms. Maryam Feily, National Advanced IPv6 Centre of Excellence (NAV6) , Universiti Sains Malaysia (USM), Malaysia

Dr. Constantine YIALOURIS, Informatics Laboratory Agricultural University of Athens, Greece

Mrs. Angeles Abella, U. de Montreal, Canada

Dr. Patrizio Arrigo, CNR ISMAC, italy

Mr. Anirban Mukhopadhyay, B.P.Poddar Institute of Management & Technology, India

Mr. Dinesh Kumar, DAV Institute of Engineering & Technology, India

Mr. Jorge L. Hernandez-Ardieta, INDRA SISTEMAS / University Carlos III of Madrid, Spain

Mr. AliReza Shahrestani, University of Malaya (UM), National Advanced IPv6 Centre of Excellence (NAv6), Malaysia

Mr. Blagoj Ristevski, Faculty of Administration and Information Systems Management - Bitola, Republic of Macedonia

Mr. Mauricio Egidio Cantão, Department of Computer Science / University of São Paulo, Brazil

Mr. Jules Ruis, Fractal Consultancy, The Netherlands

Mr. Mohammad Iftekhar Husain, University at Buffalo, USA

Dr. Deepak Laxmi Narasimha, Department of Software Engineering, Faculty of Computer Science and Information Technology, University of Malaya, Malaysia

Dr. Paola Di Maio, DMEM University of Strathclyde, UK

Dr. Bhanu Pratap Singh, Institute of Instrumentation Engineering, Kurukshetra University Kurukshetra, India

Mr. Sana Ullah, Inha University, South Korea

Mr. Cornelis Pieter Pieters, Condast, The Netherlands

Dr. Amogh Kavimandan, The MathWorks Inc., USA

Dr. Zhinan Zhou, Samsung Telecommunications America, USA

Mr. Alberto de Santos Sierra, Universidad Politécnica de Madrid, Spain

Dr. Md. Atiqur Rahman Ahad, Department of Applied Physics, Electronics & Communication Engineering (APECE), University of Dhaka, Bangladesh

Dr. Charalampos Bratsas, Lab of Medical Informatics, Medical Faculty, Aristotle University, Thessaloniki, Greece

Ms. Alexia Dini Kounoudes, Cyprus University of Technology, Cyprus

Dr. Jorge A. Ruiz-Vanoye, Universidad Juárez Autónoma de Tabasco, Mexico

Dr. Alejandro Fuentes Penna, Universidad Popular Autónoma del Estado de Puebla, México

Dr. Ocotlán Díaz-Parra, Universidad Juárez Autónoma de Tabasco, México

Mrs. Nantia Iakovidou, Aristotle University of Thessaloniki, Greece

Mr. Vinay Chopra, DAV Institute of Engineering & Technology, Jalandhar

Ms. Carmen Lastres, Universidad Politécnica de Madrid - Centre for Smart Environments, Spain

Dr. Sanja Lazarova-Molnar, United Arab Emirates University, UAE

Mr. Srikrishna Nudurumati, Imaging & Printing Group R&D Hub, Hewlett-Packard, India

Dr. Olivier Nocent, CReSTIC/SIC, University of Reims, France

Mr. Burak Cizmeci, Isik University, Turkey

Dr. Carlos Jaime Barrios Hernandez, LIG (Laboratory Of Informatics of Grenoble), France

Mr. Md. Rabiul Islam, Rajshahi university of Engineering & Technology (RUET), Bangladesh

Dr. LAKHOUA Mohamed Najeh, ISSAT - Laboratory of Analysis and Control of Systems, Tunisia

Dr. Alessandro Lavacchi, Department of Chemistry - University of Firenze, Italy

Mr. Mungwe, University of Oldenburg, Germany

Mr. Somnath Tagore, Dr D Y Patil University, India

Ms. Xueqin Wang, ATCS, USA

Dr. Borislav D Dimitrov, Department of General Practice, Royal College of Surgeons in Ireland, Dublin, Ireland

Dr. Fondjo Fotou Franklin, Langston University, USA

Dr. Vishal Goyal, Department of Computer Science, Punjabi University, Patiala, India

Mr. Thomas J. Clancy, ACM, United States

Dr. Ahmed Nabih Zaki Rashed, Dr. in Electronic Engineering, Faculty of Electronic Engineering, menouf 32951, Electronics and Electrical Communication Engineering Department, Menoufia university, EGYPT, EGYPT

Dr. Rushed Kanawati, LIPN, France

Mr. Koteswar Rao, K G Reddy College Of ENGG.&TECH,CHILKUR, RR DIST.,AP, India

Mr. M. Nagesh Kumar, Department of Electronics and Communication, J.S.S. research foundation, Mysore University, Mysore-6, India

Dr. Ibrahim Noha, Grenoble Informatics Laboratory, France

Mr. Muhammad Yasir Qadri, University of Essex, UK

Mr. Annadurai .P, KMCPGS, Lawspet, Pondicherry, India, (Aff. Pondicherry Univeristy, India

Mr. E Munivel , CEDTI (Govt. of India), India

Dr. Chitra Ganesh Desai, University of Pune, India

Mr. Syed, Analytical Services & Materials, Inc., USA

Mrs. Payal N. Raj, Veer South Gujarat University, India

Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal, India

Mr. Mahesh Goyani, S.P. University, India, India

Mr. Vinay Verma, Defence Avionics Research Establishment, DRDO, India

Dr. George A. Papakostas, Democritus University of Thrace, Greece

Mr. Abhijit Sanjiv Kulkarni, DARE, DRDO, India

Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius

Dr. B. Sivaselvan, Indian Institute of Information Technology, Design & Manufacturing, Kancheepuram, IIT Madras Campus, India

Dr. Partha Pratim Bhattacharya, Greater Kolkata College of Engineering and Management, West Bengal University of Technology, India

Mr. Manish Maheshwari, Makhanlal C University of Journalism & Communication, India

Dr. Siddhartha Kumar Khaitan, Iowa State University, USA

Dr. Mandhapati Raju, General Motors Inc, USA

Dr. M.Iqbal Saripan, Universiti Putra Malaysia, Malaysia

Mr. Ahmad Shukri Mohd Noor, University Malaysia Terengganu, Malaysia

Mr. Selvakuberan K, TATA Consultancy Services, India

Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India

Mr. Rakesh Kachroo, Tata Consultancy Services, India

Mr. Raman Kumar, National Institute of Technology, Jalandhar, Punjab., India

Mr. Nitesh Sureja, S.P.University, India

Dr. M. Emre Celebi, Louisiana State University, Shreveport, USA

Dr. Aung Kyaw Oo, Defence Services Academy, Myanmar

Mr. Sanjay P. Patel, Sankalchand Patel College of Engineering, Visnagar, Gujarat, India

Dr. Pascal Fallavollita, Queens University, Canada

Mr. Jitendra Agrawal, Rajiv Gandhi Technological University, Bhopal, MP, India

Mr. Ismael Rafael Ponce Medellín, Cenidet (Centro Nacional de Investigación y Desarrollo Tecnológico), Mexico

Mr. Shoukat Ullah, Govt. Post Graduate College Bannu, Pakistan

Dr. Vivian Augustine, Telecom Zimbabwe, Zimbabwe

Mrs. Mutalli Vatile, Offshore Business Philipines, Philipines

Mr. Pankaj Kumar, SAMA, India

Dr. Himanshu Aggarwal, Punjabi University, Patiala, India

Dr. Vauvert Guillaume, Europages, France

Prof Yee Ming Chen, Department of Industrial Engineering and Management, Yuan Ze University, Taiwan

Dr. Constantino Malagón, Nebrija University, Spain

Prof Kanwalvir Singh Dhindsa, B.B.S.B.Egg.College, Fatehgarh Sahib (Punjab), India

Mr. Angkoon Phinyomark, Prince of Singkla University, Thailand

Ms. Nital H. Mistry, Veer Narmad South Gujarat University, Surat, India

Dr. M.R.Sumalatha, Anna University, India

Mr. Somesh Kumar Dewangan, Disha Institute of Management and Technology, India

Mr. Raman Maini, Punjabi University, Patiala(Punjab)-147002, India

Dr. Abdelkader Outtagarts, Alcatel-Lucent Bell-Labs, France

Prof Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India

Mr. Prabu Mohandas, Anna University/Adhiyamaan College of Engineering, india

Dr. Manish Kumar Jindal, Panjab University Regional Centre, Muksar, India

Prof Mydhili K Nair, M S Ramaiah Institute of Technnology, Bangalore, India

Dr. C. Suresh Gnana Dhas, VelTech MultiTech Dr.Rangarajan Dr.Sagunthala Engineering College,Chennai,Tamilnadu, India

Prof Akash Rajak, Krishna Institute of Engineering and Technology, Ghaziabad, India

Mr. Ajay Kumar Shrivastava, Krishna Institute of Engineering & Technology, Ghaziabad, India

Dr. Vu Thanh Nguyen, University of Information Technology HoChiMinh City, VietNam

Prof Deo Prakash, SMVD University (A Technical University open on I.I.T. Pattern) Kakryal (J&K), India

Dr. Navneet Agrawal, Dept. of ECE, College of Technology & Engineering, MPUAT, Udaipur 313001 Rajasthan, India

Mr. Sufal Das, Sikkim Manipal Institute of Technology, India

Mr. Anil Kumar, Sikkim Manipal Institute of Technology, India

Dr. B. Prasanalakshmi, King Saud University, Saudi Arabia.

Dr. K D Verma, S.V. (P.G.) College, Aligarh, India

Mr. Mohd Nazri Ismail, System and Networking Department, University of Kuala Lumpur (UniKL), Malaysia

Dr. Nguyen Tuan Dang, University of Information Technology, Vietnam National University Ho Chi Minh city, Vietnam

Dr. Abdul Aziz, University of Central Punjab, Pakistan

Dr. P. Vasudeva Reddy, Andhra University, India

Mrs. Savvas A. Chatzichristofis, Democritus University of Thrace, Greece

Mr. Marcio Dorn, Federal University of Rio Grande do Sul - UFRGS Institute of Informatics, Brazil

Mr. Luca Mazzola, University of Lugano, Switzerland

Mr. Hafeez Ullah Amin, Kohat University of Science & Technology, Pakistan

Dr. Professor Vikram Singh, Ch. Devi Lal University, Sirsa (Haryana), India

Dr. Shahanawaj Ahamad, Department of Computer Science, King Saud University, Saudi Arabia

Dr. K. Duraiswamy, K. S. Rangasamy College of Technology, India

Prof. Dr Mazlina Esa, Universiti Teknologi Malaysia, Malaysia

Dr. P. Vasant, Power Control Optimization (Global), Malaysia

Dr. Taner Tuncer, Firat University, Turkey

Dr. Norrozila Sulaiman, University Malaysia Pahang, Malaysia

Prof. S K Gupta, BCET, Guradspur, India

Dr. Latha Parameswaran, Amrita Vishwa Vidyapeetham, India

Mr. M. Azath, Anna University, India

Dr. P. Suresh Varma, Adikavi Nannaya University, India

Prof. V. N. Kamalesh, JSS Academy of Technical Education, India

Dr. D Gunaseelan, Ibri College of Technology, Oman

Mr. Sanjay Kumar Anand, CDAC, India

Mr. Akshat Verma, CDAC, India

Mrs. Fazeela Tunnisa, Najran University, Kingdom of Saudi Arabia

Mr. Hasan Asil, Islamic Azad University Tabriz Branch (Azarshahr), Iran

Prof. Dr Sajal Kabiraj, Fr. C Rodrigues Institute of Management Studies (Affiliated to University of Mumbai, India), India

Mr. Syed Fawad Mustafa, GAC Center, Shandong University, China

Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA

Prof. Selvakani Kandeegan, Francis Xavier Engineering College, India

Mr. Tohid Sedghi, Urmia University, Iran

Dr. S. Sasikumar, PSNA College of Engg and Tech, Dindigul, India

Dr. Anupam Shukla, Indian Institute of Information Technology and Management Gwalior, India

Mr. Rahul Kala, Indian Institute of Information Technology and Management Gwalior, India

Dr. A V Nikolov, National University of Lesotho, Lesotho

Mr. Kamal Sarkar, Department of Computer Science and Engineering, Jadavpur University, India

Prof. Sattar J Aboud, Iraqi Council of Representatives, Iraq-Baghdad

Dr. Prasant Kumar Pattnaik, Department of CSE, KIST, India

Dr. Mohammed Amoon, King Saud University, Saudi Arabia

Dr. Tsvetanka Georgieva, Department of Information Technologies, St. Cyril and St. Methodius University of Veliko Tarnovo, Bulgaria

Mr. Ujjal Marjit, University of Kalyani, West-Bengal, India

Dr. Prasant Kumar Pattnaik, KIST, Bhubaneswar, India, India

Dr. Guezouri Mustapha, Department of Electronics, Faculty of Electrical Engineering, University of Science and Technology (USTO), Oran, Algeria

Mr. Maniyar Shiraz Ahmed, Najran University, Najran, Saudi Arabia

Dr. Sreedhar Reddy, JNTU, SSIETW, Hyderabad, India

Mr. Bala Dhandayuthapani Veerasamy, Mekelle University, Ethiopia

Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia

Mr. Rajesh Prasad, LDC Institute of Technical Studies, Allahabad, India

Ms. Habib Izadkhah, Tabriz University, Iran

Dr. Lokesh Kumar Sharma, Chhattisgarh Swami Vivekanand Technical University Bhilai, India

Mr. Kuldeep Yadav, IIT Delhi, India

Dr. Naoufel Kraiem, Institut Supérieur d'Informatique, Tunisia

Prof. Frank Ortmeier, Otto-von-Guericke-Universität Magdeburg, Germany

Mr. Ashraf Aljammal, USM, Malaysia

Mrs. Amandeep Kaur, Department of Computer Science, Punjabi University, Patiala, Punjab, India

Mr. Babak Basharirad, University Technology of Malaysia, Malaysia

Mr. Avinash Singh, Kiet Ghaziabad, India

Dr. Miguel Vargas-Lombardo, Technological University of Panama, Panama

Dr. Tuncay Sevindik, Firat University, Turkey

Ms. Pavai Kandavelu, Anna University Chennai, India

Mr. Ravish Khichar, Global Institute of Technology, India

Mr. Ahsan Ali Zaidan Ansaef, Multimedia University, Cyberjaya, Malaysia

Dr. Awadhesh Kumar Sharma, Dept. of CSE, MMM Engg College, Gorakhpur-273010, UP, India

Mr. Qasim Siddique, FUIEMS, Pakistan

Dr. Le Hoang Thai, University of Science, Vietnam National University - Ho Chi Minh City, Vietnam

Dr. Saravanan C, NIT, Durgapur, India

Dr. Vijay Kumar Mago, DAV College, Jalandhar, India

Dr. Do Van Nhon, University of Information Technology, Vietnam

Dr. Georgios Kioumourtzis, Researcher, University of Patras, Greece

Mr. Amol D. Potgantwar, SITRC Nasik, India

Mr. Lesedi Melton Masisi, Council for Scientific and Industrial Research, South Africa

Dr. Karthik.S, Department of Computer Science & Engineering, SNS College of Technology, India

Mr. Nafiz Imtiaz Bin Hamid, Department of Electrical and Electronic Engineering, Islamic University of Technology (IUT), Bangladesh

Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia

Dr. Abdul Kareem M. Radhi, Information Engineering - Nahrain University, Iraq

Dr. Manuj Darbari, BBDNITM, Institute of Technology, A-649, Indira Nagar, Lucknow 226016, India

Ms. Izerrouken, INP-IRIT, France

Mr. Nitin Ashokrao Naik, Dept. of Computer Science, Yeshwant Mahavidyalaya, Nanded, India

Mr. Nikhil Raj, National Institute of Technology, Kurukshetra, India

Prof. Maher Ben Jemaa, National School of Engineers of Sfax, Tunisia

Prof. Rajeshwar Singh, BRCM College of Engineering and Technology, Bahal Bhiwani, Haryana, India

Mr. Gaurav Kumar, Department of Computer Applications, Chitkara Institute of Engineering and Technology, Rajpura, Punjab, India

Mr. Ajeet Kumar Pandey, Indian Institute of Technology, Kharagpur, India

Mr. Rajiv Phougat, IBM Corporation, USA

Mrs. Aysha V, College of Applied Science Pattuvam affiliated with Kannur University, India

Dr. Debotosh Bhattacharjee, Department of Computer Science and Engineering, Jadavpur University, Kolkata-700032, India

Dr. Neelam Srivastava, Institute of engineering & Technology, Lucknow, India

Prof. Sweta Verma, Galgotia's College of Engineering & Technology, Greater Noida, India

Mr. Harminder Singh Bindra, MIMIT, INDIA

Mr. Tarun Kumar, U.P. Technical University/Radha Govinend Engg. College, India

Mr. Tirthraj Rai, Jawahar Lal Nehru University, New Delhi, India

Mr. Akhilesh Tiwari, Madhav Institute of Technology & Science, India

Mr. Dakshina Ranjan Kisku, Dr. B. C. Roy Engineering College, WBUT, India

Ms. Anu Suneja, Maharshi Markandeshwar University, Mullana, Haryana, India

Mr. Munish Kumar Jindal, Punjabi University Regional Centre, Jaito (Faridkot), India

Dr. Ashraf Bany Mohammed, Management Information Systems Department, Faculty of Administrative and Financial Sciences, Petra University, Jordan

Mrs. Jyoti Jain, R.G.P.V. Bhopal, India

Dr. Lamia Chaari, SFAX University, Tunisia

Mr. Akhter Raza Syed, Department of Computer Science, University of Karachi, Pakistan

Prof. Khubaib Ahmed Qureshi, Information Technology Department, HIMS, Hamdard University, Pakistan

Prof. Boubker Sbihi, Ecole des Sciences de L'Information, Morocco

Dr. S. M. Riazul Islam, Inha University, South Korea

Prof. Lokhande S.N., S.R.T.M.University, Nanded (MH), India

Dr. Vijay H Mankar, Dept. of Electronics, Govt. Polytechnic, Nagpur, India

Mr. Ojesanmi Olusegun, Ajayi Crowther University, Oyo, Nigeria

Ms. Mamta Juneja, RBIEBT, PTU, India

Prof. Chandra Mohan, John Bosco Engineering College, India

Dr. Bodhe Shrikant K., College of Engineering, Pandhapur, Maharashtra, INDIA

Dr. Sherif G. Aly, The American University in Cairo, Egypt

Mr. Sunil Kashibarao Nayak, Bahirji Smarak Mahavidyalaya, Basmathnagar Dist-Hingoli., India

Prof. Nikhil gondaliya, G H Patel College of Engg. & Technology, India

Mr. Nisheeth Joshi, Apaji Institute, Banasthali University, India

Mr. Nizar, National Engineering School of Monastir, Tunisia

Prof. R. Jagadeesh Kannan, RMK Engineering College, India

Prof. Rakesh.L, Vijetha Institute of Technology, Bangalore, India

Mr B. M. Patil, Indian Institute of Technology, Roorkee, Uttarakhand, India

Dr. Intisar A. M. Al Sayed, Associate prof./College of Science and IT/Al Isra University, Jordan

Mr. Thipendra Pal Singh, Sharda University, K.P. III, Greater Noida, Uttar Pradesh, India

Mrs. Rajalakshmi, JIITU, India

Mr. Shrikant Ardhapurkar, Indian Institute of Information Techonology, India

Ms. Hemalatha R, Osmania University, India

Mr. Hadi Saboohi, University of Malaya - Faculty of Computer Science and Information Technology, Malaysia

Mr. Sunil Kumar Grandhi, Maris Stella College, India

Prof. Shishir K. Shandilya, NRI Institute of Science & Technology, INDIA

Dr. Umesh Kumar Singh, Vikram University, Ujjain, India

Prof. Prasun Ghosal, Bengal Engineering and Science University, India

Dr. Nagarajan Velmurugan, SMVEC/Pondicherry University, India

Dr. R. Baskaran, Anna University, India

Dr. Wichian Sittiprapaporn, Mahasarakham University College of Music, Thailand

Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia

Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology, India

Mrs. Inderpreet Kaur, PTU, Jalandhar, India

Mr. Palaniyappan, K7 Virus Research Laboratory, India

Mr. Guanbo Zheng, University of Houston, main campus, USA

Mr. Arun Kumar Tripathi, Krishna Institute of Engg. and Tech-Ghaziabad, Affiliated to UPTU, India

Mr. Iqbaldeep Kaur, PTU / RBIEBT, India

Mr. Amit Choudhary, Maharaja Surajmal Institute, New Delhi, India

Mrs. Vasudha Bahl, Maharaja Agrasen Institute of Technology, Delhi, India

Dr. Ashish Avasthi, Uttar Pradesh Technical University, India

Dr. Manish Kumar, Uttar Pradesh Technical University, India

Prof. Vinay Uttamrao Kale, P.R.M. Institute of Technology & Research, Badnera, Amravati, Maharashtra, India

Mr. Suhas J Manangi, Microsoft, India

Mr. Shyamalendu Kandar, Haldia Institute of Technology, India

Ms. Anna Kuzio, Adam Mickiewicz University, School of English, Poland

Mr. Vikas Singla, Malout Institute of Management & Information Technology, Malout, Punjab, India, India

Dr. Dalbir Singh, Faculty of Information Science And Technology, National University of Malaysia, Malaysia

Dr. Saurabh Mukherjee, PIM, Jiwaji University, Gwalior, M.P, India

Mr. Senthilnathan T, Sri Krishna College of Engineering and Technology, India

Dr. Debojyoti Mitra, Sir Padampat Singhania University, India

Prof. Rachit Garg, Department of Computer Science, L K College, India

Dr. Arun Kumar Gupta, M.S. College, Saharanpur, India

Dr. Todor Todorov, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Bulgaria

Mrs. Manjula K A, Kannur University, India

Mrs. Sasikala R., K S R College of Technology, India

Prof. M. Saleem Babu, Department of Computer Science and Engineering, Vel Tech University, Chennai, India

Dr. Rajesh Kumar Tiwari, GLA Institute of Technology, India

Mr. Rakesh Kumar, Indian Institute of Technology Roorkee, India

Prof. Amit Verma, PTU/RBIEBT, India

Mr. Sohan Purohit, University of Massachusetts Lowell, USA

Mr. Anand Kumar, AMC Engineering College, Bangalore, India

Dr. Samir Abdelrahman, Computer Science Department, Cairo University, Egypt

Dr. Rama Prasad V Vaddella, Sree Vidyanikethan Engineering College, India

Dr. Manoj Wadhwa, Echelon Institute of Technology Faridabad, India

Mr. Zeashan Hameed Khan, Universit  de Grenoble, France

Mr. Arup Kumar Pal, Indian School of Mines, Dhanbad, India

Dr. Pouya, Islamic Azad University, Naein Branch, Iran

Prof. Jyoti Prakash Singh, Academy of Technology, India

Mr. Muraleedharan CV, Sree Chitra Tirunal Institute for Medical Sciences & Technology, India

Dr. E U Okike, University of Ibadan, Nigeria Kampala Int Univ Uganda, Nigeria

Dr. D. S. Rao, Chitkara University, India

Mr. Peyman Taher, Oklahoma State University, USA

Dr. S Srinivasan, PDM College of Engineering, India

Dr. Rafiqul Zaman Khan, Department of Computer Science, AMU, Aligarh, India

Ms. Meenakshi Kalia, Shobhit University, India

Mr. Muhammad Zakarya, Abdul Wali Khan University, Mardan, Pakistan, Pakistan

Dr. M Gobi, PSG college, India

Mr. Williamjeet Singh, Chitkara Institute of Engineering and Technology, India

Mr. G.Jeyakumar, Amrita School of Engineering, India

Mr. Osama Sohaib, University of Balochistan, Pakistan

Mr. Jude Hemanth, Karunya University, India

Mr. Nitin Rakesh, Jaypee University of Information Technology, India

Mr. Harmunish Taneja, Maharishi Markandeshwar University, Mullana, Ambala, Haryana, India

Dr. Sin-Ban Ho, Faculty of IT, Multimedia University, Malaysia

Dr. Mashiur Rahman, Institute for Molecular Science, Japan

Mrs. Doreen Hephzibah Miriam, Anna University, Chennai, India

Mr. Kosala Yapa Bandara, Dublin City University, Ireland.

Mrs. Mitu Dhull, GNKITMS Yamuna Nagar Haryana, India

Dr. Chitra A.Dhawale, Professor, Symbiosis Institute of Computer Studies and Research, Pune (MS), India

Dr. Arun Sharma, GB Technical University, Noida, India

Mr. Naoufel Machta, Faculty of Science of Tunis, Tunisia

Dr. Utpal Biswas, University of Kalyani, India

Prof. Parma Nand, IIT Roorkee, India

Prof. Mahesh P K, Jnana Vikas Institute of Tevhnology, Bangalore, India

Dr. D.I. George Amalarethnam, Jamal Mohamed College, Bharathidasan University, India

Mr. Ishtiaq ahmad, University of Engineering & Technology, Taxila, Pakistan

Mrs. B.Sharmila, Sri Ramakrishna Engineering College, Coimbatore Anna University Coimbatore, India

Dr. Muhammad Wasif Nisar, COMSATS Institue of Information Technology, Pakistan

Mr. Prabu Dorairaj, EMC Corporation, India/USA

Mr. Neetesh Gupta, Technocrats Inst. of Technology, Bhopal, India

Dr. Ola Osunkoya, PRGX, USA

Ms. A. Lavanya, Manipal University, Karnataka, India

Dr. Jalal Laassiri, MIA-Laboratory, Faculty of Sciences Rabat, Morocco

Mr. Ganesan, Sri Venkateswara college of Engineering and Technology, Thiruvallur, India

Mr. V.Ramakrishnan, Sri Venkateswara college of Engineering and Technology, Thiruvallur, India

Prof. Vuda Sreenivasarao, St. Mary's college of Engg & Tech, India

Prof. Ashutosh Kumar Dubey, Assistant Professor, India

Dr. R.Ramesh, Anna University, India

Mr. Ali Khadair HMood, University of Malaya, Malaysia

Dr. Vimal Mishra, U.P. Technical Education, India

Mr. Ranjit Singh, Apeejay Institute of Management, Jalandhar, India

Mrs. D.Suganyadevi, SNR SONS College (Autonomous), India

Mr. Prasad S.Halgaonkar, MIT, Pune University, India

Mr. Vijay Kumar, College of Engg. and Technology, IFTM, Moradabad(U.P), India

Mr. Mehran Parchebafieh, Douran, Iran

Mr. Anand Sharma, MITS, Lakshmanagarh, Sikar (Rajasthan), India

Mr. Amit Kumar, Jaypee University of Engineering and Technology, India

Prof. B.L.Shivakumar, SNR Sons College, Coimbatore, India

Mr. Mohammed Imran, JMI, India

Dr. R Bremananth, School of EEE, Information Engineering (Div.), Nanyang Technological University, Singapore

Prof. Vasavi Bande, Computer Science and Engineering, Hyderabad Institute of Technology and Management, India

Dr. S.R.Balasundaram, National Institute of Technology, India

Dr. Prasart Nuangchalerm, Mahasarakham University, Thailand

Dr. M Ayoub Khan, C-DAC, Ministry of Communications & IT., India

Dr. Jagdish Lal Raheja, Central Electronics Engineering Research Institute, India

Mr G. Appasami, Dept. of CSE, Dr. Pauls Engineering College, Anna University - Chennai, India

Mr Vimal Mishra, U.P. Technical Education, Allahabad, India

Mr. Amin Daneshmand Malayeri, Young Researchers Club, Islamic AZAD University, Malayer Branch, Iran

Dr. Arti Arya, PES School of Engineering, Bangalore (under VTU, Belgaum, Karnataka), India

Mr. Pawan Jindal, J.U.E.T. Guna, M.P., India

Dr. Soumen Mukherjee, RCC Institute of Information Technology, India

Dr. Hamid Mcheick, University of Qubec at Chicoutimi, Canada

Dr. Mokhled AlTarawneh, PhD computer engineering/ Faculty of engineering/ mutah university, Jordan

Prof. Santhosh.P.Mathew, Saintgits College of Engineering, Kottayam, India

Ms. Suman Lata, Rayat Bahara institute of engg. & Nanotechnology, Hoshiarpur, India

Dr. Shaikh Abdul Hannan, Vivekanand College, Aurangabad, India

Prof. PN Kumar, Amrita Vishwa Vidyapeetham, India

Dr. P. K. Suri, Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, India

Dr. Syed Akhter Hossain, Daffodil International University, Bangladesh

Mr. Sunil, Vignan College, India

Mr. Ajit Singh, TIT&S Bhiwani, Haryana, India

Mr. Nasim Qaisar, Federal Urdu University of Arts, Science and Technology, Pakistan

Ms. Rshma, Maharishi Markandeshwar University, India

Mr. Gaurav Kumar Leekha, M.M.University, Solan (Himachal Pradesh), India

Mr. Ordinar Tucker, Ministry of Finance Jamaica, Jamaica

Mr. Mohit Jain, Maharaja Surajmal Institute of Technology (Affiliated to Guru Gobind Singh Indraprastha University, New Delhi), India

Dr. Shaveta Rani, GZS College of Engineering & Technology, India

Dr. Paramjeet Singh, GZS College of Engineering & Technology, India

Dr. G R Sinha, SSCET, India

Mr. Chetan Sharma, TechMahindra India Ltd., India

Dr. Nabil Mohammed Ali Munassar, University of Science and Technology, Yemen

Prof. T Venkat Narayana Rao, Department of CSE, Hyderabad Institute of Technology and Management, India

Prof. Vasavi Bande, HITAM, Engineering College, India

Prof. S.P.Setty, Andhra University, India

Dr. C. Kiran Mai, J.N.T.University, Hyderabad/VNR Vignana Jyothi Institute of Engineering & Technology/, India

Ms. Bindiya Ahuja, Manav Rachna International University, India

Mrs. Deepa Bura, Manav Rachna International University, India

Mr. Vikas Gupta, CDLM Government Engineering College, Panniwala Mota, India

Dr Juan José Martínez Castillo, University of Yacambu, Venezuela

Mr Kunwar S. Vaisla, Department of Computer Science & Engineering, BCT Kumaon Engineering College, India

Mr. Abhishek Shukla, RKGIT, India

Prof. Manpreet Singh, M. M. Engg. College, M. M. University, Haryana, India

Mr. Syed Imran, University College Cork, Ireland

Dr. Intisar Al Said, Associate Prof/Al Isra University, Jordan

Dr. Namfon Assawamekin, University of the Thai Chamber of Commerce, Thailand

Dr. Shiv KUMar, Technocrat Institute of Technology-Bhopal (M.P.), India

Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran

Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran

Dr. Mohamed Ali Mahjoub, University of Monastir, Tunisia

Mr. Adis Medic, Infosys Ltd, Bosnia and Herzegovina

Mr Swarup Roy, Department of Information Technology, North Eastern Hill University, Umshing, Shillong 793022, Meghalaya, India

Prof. Jakimi, Faculty of Science and technology my ismail University, Morocco

Dr. R. Manicka Chezian, N G M College, Pollachi - 642 001, Tamilnadu, India

Dr. P.Dananjayan, Pondicherry Engineering College, India

Mr. Manik Sharma, Sewa Devi SD College Tarn Taran, India

Mr. Suresh Kallam, East China University of Technology, Nanchang, China

Dr. Mohammed Ali Hussain, Sai Madhavi Institute of Science & Technology, Rajahmundry, India

Mr. Vikas Gupta, Adesh Institute of Engineering & Technology, India

Dr. Anuraag Awasthi, JV Womens University, Jaipur, India

Dr. Mathura Prasad Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), Srinagar (Garhwal), India

Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia, Malaysia

Mr. Adnan Qureshi, University of Jinan, Shandong, P.R.China, P.R.China

Dr. Jatinderkumar R. Saini, Narmada College of Computer Application, India

Mr. Mueen Uddin, Universiti Teknologi Malaysia, Malaysia

Mr. Manoj Gupta, Apex Institute of Engineering & Technology, Jaipur (Affiliated to Rajasthan Technical University, Rajasthan), Indian

Mr. S. Albert Alexander, Kongu Engineering College, India

Dr. Shaidah Jusoh, Zarqa Private University, Jordan

Dr. Dushmanta Mallick, KMBB College of Engineering and Technology, India

Mr. Santhosh Krishna B.V, Hindustan University, India

Dr. Tariq Ahamad Ahanger, Kausar College Of Computer Sciences, India

Dr. Chi Lin, Dalian University of Technology, China

Prof. VIJENDRA BABU.D, ECE Department, Aarupadai Veedu Institute of Technology, Vinayaka Missions University, India

Mr. Raj Gaurang Tiwari, Gautam Budh Technical University, India

Mrs. Jeysree J, SRM University, India

Dr. C S Reddy, VIT University, India

Dr. Amit Wason, Rayat-Bahra Institute of Engineering & Bio-Technology, Kharar, India

Mr. Muhammad Shuaib Qureshi, Iqra National University, Peshawar, Pakistan, Pakistan

Dr Pranam Paul, Narula Institute of Technology Agarpara. Kolkata: 700109; West Bengal, India

Dr. G. M. Nasira, Sasurie College of Engineering, (Affiliated to Anna University of Technology Coimbatore), India

Dr. Manasawee Kaenampornpan, Mahasarakham University, Thailand

Mrs. Iti Mathur, Banasthali University, India

Mr. Avanish Kumar Singh, RRIMT, NH-24, B.K.T., Lucknow, U.P., India

Mr. Velayutham Pavanasam, Adhiparasakthi Engineering College, Melmaruvathur, India

Dr. Panagiotis Michailidis, University of Western Macedonia, Greece

Mr. Amir Seyed Danesh, University of Malaya, Malaysia

Dr. Nadeem Mahmood, Department of computer science, university of Karachi, Pakistan

Dr. Terry Walcott, E-Promag Consultancy Group, United Kingdom

Mr. Farhat Amine, High Institute of Management of Tunis, Tunisia

Mr. Ali Waqar Azim, COMSATS Institute of Information Technology, Pakistan

Mr. Zeeshan Qamar, COMSATS Institute of Information Technology, Pakistan

Dr. Samsudin Wahab, MARA University of Technology, Malaysia

Mr. Ashikali M. Hasan, CelNet Security, India

Dr. Binod Kumar, Lakshmi Narayan College of Tech.(LNCT), India

Mr. B V A N S S Prabhakar Rao, Dept. of CSE, Miracle Educational Society Group of Institutions, Vizianagaram, India

Dr. T. Abdul Razak, Associate Professor of Computer Science, Jamal Mohamed College (Affiliated to Bharathidasan University, Tiruchirappalli), Tiruchirappalli-620020, India

Mr. Aurobindo Ogra, University of Johannesburg, South Africa

Mr. Essam Halim Houssein, Dept of CS - Faculty of Computers and Informatics, Benha - Egypt

Dr. Hanumanthappa. J, DoS in Computer Science, India

Mr. Rachit Mohan Garg, Jaypee University of Information Technology, India

Mr. Kamal Kad, Infosys Technologies, Australia

Mrs. Aditi Chawla, GNIT Group of Institutes, India

Dr. Kumardatt Ganrje, Pune University, India

Mr. Merugu Gopichand, JNTU/BVRIT, India

Mr. Rakesh Kumar, M.M. University, Mullana,Ambala, India

Mr. M. Sundar, IBM, India

Prof. Mayank Singh, J.P. Institute of Engineering & Technology, India

Dr. Saurabh Pal, VBS Purvanchal University, Jaunpur, India

Mr. Khaleel Ahmad, S.V.S. University, India

Mr. Amin Zehtabian, Babol Noshirvani University of Technology / Tetta Electronic Company, Iran

Mr. Rahul Katarya, Department of Information Technology , Delhi Technological University, India

Dr. Vincent Ele Asor, University of Port Harcourt, Nigeria

Ms. Prayas Kad, Capgemini Australia Ltd, Australia

Mr. Alireza Jolfaei, Faculty and Research Center of Communication and Information Technology, IHU, Iran

Mr. Nitish Gupta, GGSIPU, India

Dr. Mohd Lazim Abdullah, University of Malaysia Terengganu, Malaysia

Ms. Suneet Kumar, Uttarakhand Technical University/Dehradun Institute of Technology, Dehradun, Uttarakhand, India

Mr. Rupesh Nasre., Indian Institute of Science, Bangalore., India.

Mrs. Dimpi Srivastava, Dept of Computer science, Information Technology and Computer Application, MIET, Meerut, India

Dr. Eva Volna, University of Ostrava, Czech Republic

Prof. Santosh Balkrishna Patil, S.S.G.M. College of Engineering, Shegaon, India

Mr. Mohd Dilshad Ansari, Jaypee University of Information Technology Solan (HP), India

Mr. Ashwani Kumar, Jaypee University of Information Technology Solan(HP), India

Dr. Abbas Karimi, Faculty of Engineering, I.A.U. Arak Branch, Iran

Mr. Fahimuddin.Shaik, AITS, Rajampet, India

Mr. Vahid Majid Nezhad, Islamic Azad University, Iran

Ms. C. Divya, Dr G R Damodaran College of Science, Coimbatore-641014, Tamilnadu, India

Prof. D. P. Sharma, AMU, Ethiopia

Dr. Sukumar Senthilkumar, School of Mathematical Sciences, Universiti Sains Malaysia, Malaysia

Mr. Sanjay Bhargava, Banasthali University, Jaipur, Rajasthan, India

Prof. Rajesh Deshmukh, Shri Shankaracharya Institute of Professional Management & Technology, India

Mr. Shervan Fekri Ershad, shiraz international university, Iran

Dr. Vladimir Urosevic, Ministry of Interior, Republic of Serbia

Mr. Ajit Singh, MDU Rohtak, India

Prof. Asha Ambhaikar, Rungta College of Engineering & Technology, Bilai, India

Dr. Saurabh Dutta, Dr. B. C. Roy Engineering College, Durgapur, India

Dr. Mokhled Altarawneh, Mutah University, Jordan

Mr. Anand Nayyar, KCL Institute of Management and Technology, Jalandhar, India

Mr S. A. Ahsan rajon, Computer Science and Engineering Discipline, Khulna University, Bangladesh

Ms. Rezarta Mersini, University of Durres, Albania

Mrs. Deepika Joshi, Jaipuria Institute of Management Studies, India

Dr. Niraj Shakhakarmi, Prairie View A&M University, (Texas A&M University System), USA

Mrs. A. Valarmathi, Anna University, Trichy, India

Dr. K. Balamurugan, Institute of Road and Transport Technology, India

Prof. K S Sridharan, Sri Sathya Sai Institute of Higher Learning, India

Mr. Okumoku-Evrero Oniovosa, Delta State University, Abraka, Nigeria

Mr. Rajiv Chopra, GTBIT, Delhi, India

Mr. Harish Garg, Department of Mathematics, IIT Roorkee, India

Mr. Ganesh Davanam, Sree Vidyanikethan Engineering College, India

Mr. Bhavesh Shah, VIT, India

Dr. Suresh Kumar Bhardwaj, Manav Rachna International University, India

Dr. Muhammad Nawaz Khan, School of electrical engineering & Computer Science, Pakistan

Ms. Saranya, Bharathidasan University, India

Mr. Sumit Joshi, GRD-IMT, Dehradun, India

Dr. Mohammed M. Abu Shquier, Tabuk University, School of Computers and Information Technology, Kingdom of Saudi Arabia

Ms. Shalini Ramanathan, PSG College of Technology, India

Mr. S.Munisankaraiah, Geethanjali college of Engineering & Technology, Hyderabad, India

Dr. Satyanarayana, KL University, India

Mr. Sarin CR, Anna University, India

Mr. Sayed Shoaib Anwar, Mahatma Gandhi Mission College of Engineering, India

Mrs. Gunjan, JSSATE, Noida, India

Dr. Ramachandra V Pujeri, Anna University, India

Mrs. Antima Singh Puniya, Shobhit University, Meerut, India

Dr. Avdhesh Gupta, College of Engineering Roorkee, India

Ms. Shiva Prakash, Madan Mohan Malaviya Engg. College, Gorakhpur, India

Dr. Kristijan Kuk, School of Electrical Engineering and Computer Science Applied Studies, Belgrade, Serbia

Prof. Dinesh Vitthalrao Rojatar, Govt. College of Engineering, Chandrapur, India

Prof. Lalji Prasad, RGTU/TCET, Indore, india

Dr. A. John Sanjeev Kumar, Thiagarajar College of Engineering, Madurai, Tamilnadu, India

Mr. Harishbabu Kalidasu, Priyadarshini Institute of Technology and Science, Tenali, Guntur(DT), Andhra Pradesh, India

Prof. Vaitheeshwaran, Priyadarshini Indira Gandhi College of Engineering, India

Mrs. P.Salini, Pondicherry Engineering College, India

Mr. Vivek Bhambri, Desh Bhagat Institute of Management and Computer Sciences, Mandi Gobindgarh(Punjab), India

Mr. Slavko Zitnik, Faculty of Computer and Information Science Ljubljana, Slovenia

Ms. Sreenivasa Rao, CMJ University/Yodlee Infotech, India

Mr. Shihabudheen P M, TATA ELXSI LTD, India

Dr. Ahmed Moustafa Elmahalawy, Faculty of Electronics Engineering, Computer Science and Engineering, Egypt

Mr. Kamlesh Kumar, Kumaun University, Nainital, India

TABLE OF CONTENTS

1. The Design of a Dual-band antenna for GPS-Beidou Applications Shenzheng Zuo, Luyao Shi and Encheng Wang	1-3
2. Architecture Exploration of Multicore Systems-on-Chip using a TLM-based Framework Mona Safar, Magdy A. El-Moursy, Mohamed Abdelsalam and Ashraf Salem	4-7
3. Three-dimensional Numerical Study of Natural Convection in a Cubical Enclosure with Two Heated Square Sections Submitted to Periodic Temperatures Lahoucine Belarche, Btissam Abourida and Touria Mediouni	8-16
4. Characteristics of Global Calling in VoIP services: A logistic regression analysis Thanyalak Mueangkhrot, Julian Ming-Sung Cheng and Chaknarin Kongcharoen	17-22
5. EVAM-MAC: An Event Based Medium Access Control for Wireless sensor Networks with Multihop Support Zaher Merhi, Mohamed Elgamel, Samih Abdul-Nabi, Amin Haj-Ali and Magdy Bayoumi	23-38
6. A Critical Review of Development of Intelligent Tutoring Systems: Retrospect, Present and Prospect Neelu Jyothi Ahuja and Roohi Sille	39-48
7. Mitigate Black Hole Attack In Dynamic Source Routing (DSR) Protocol By Trapping Mahamuni K and Chandrasekar C	49-54
8. Transformation from E-commerce to M-commerce in Indian Context Sujata P. Deshmukh, Prashant Deshmukh and G.T. Thampi	55-60
9. An Expressive Role-based Approach for Improving Distributed Collaboration Transparency Abdullah O. Al-Zaghameem and Mohammad Alfraheed	61-67
10. Exit Chart for Iterative Decoding of Product and Concatenated Block Codes Abderrazak Farchane, Mostafa Belkasmi and Ahmed Azouaoui	68-77
11. Protection of Personal Data in Information Systems Bojken Shehu, Shqiponja Ahmetaj, Marin Aranitasi and Aleksander Xhuvani	78-81
12. The Analysis and Design of a Web-Based Social Network: Locate it! Project Noura Alomar and Wea'am Alrashed	82-88
13. An Eccentric Scheme for Oblivious Communication Khan Farhan Rafat and Muhammad Sher	89-96
14. Performance analysis of Internet Protocol Storage Area Network (IP SAN) and its usage in Clustered Database Akanksha Verma, Shrijee Bhattacharjee, Prashant Baste, Nandan Nagarajappa Mirajkar, Sandeep Bhujbal, Nilesh Deshmukh and Kumud Singh	97-104

15. Towards an Assessment-oriented Model for External Information System Quality Characterization Abir Elmir, Badr Elmir and Bouchaib Bounabat	105-112
16. Towards a better control of the infectious risks associated with health care activities Benzouai Messaoud, Smadi Hacene, Kassah-Laouar Ahmed, Bahmani Younes and Bellala Djamel	113-121
17. Image enhancement using fusion by wavelet transform and Laplacian pyramid Shailendra M. Mukane, Yogesh S Ghodake and P.S.Khandagle	122-126
18. User Authentication with Adaptive Keystroke Dynamics Shimaa I. Hassan, Mazen M. Selim and Hala H. Zayed	127-134
19. Indiscernible Communication through ASCII Text Document/File (Communication in Veil) Khanfarhanrafat and Muhammad Sher	135-140
20. Biometric System by Pointing Finger Eugène Mbuyi Mukendi and Jean Didier Batubenga Muamba	141-149
21. Sapwood Oak classification using Global and Local Features Hammad Naeem, Asim Anees, Maria Minhas and Jameel Ahmed	150-153
22. A Comparative Study about Object Classification based on Global and Local Features Hammad Naeem, Maria Minhas and Jameel Ahmed	154-158
23. Cloud Computing: A Survey on its limitations and Potential Solutions Mohammad Manzurul Islam, Sarwar Morshed and Parijat Goswami	159-163
24. Intelligent Vehicle Recognition based on Wireless Sensor Network Maha Mohamed Nabeel, Mahmoud Fakhr El-Dein and Sherine Abd El-Kader	164-174
25. Online Newspaper using Mashup Mentor Hamiti and Ramiz Kastrati	175-179
26. Lock-Free Readers/Writers Anupriya Chakraborty, Sourav Saha, Ryan Saptarshi Ray and Utpal Kumar Ray	180-186
27. The Utilization of Internet Communication into Knowledge Sharing Activities in Primary Schools Lokman Mohd Tahir, Mohd Fadzli Ali, Tina Abdullah, Mehmet Ozay, Haswadi Hassan and Khadijah Daud	187-194
28. Adoption of ICT-Based Secondary Education for Actualizing the Millenium Development Goals in Nigeria O.A. Odeniyi, A.Q Ayinde, N.B Lawal, A.B Ghazaly-Agboola and S.O Kolade	195-199

29. Minimizing Handoff Latency and Packet Loss in NEMO Fatimah Abdulnabi Salman and Emad Hassan Al-Hemairy	200-207
30. A Proposed Permutation Scheme based on 3-D Chaotic System for Encrypting the Colored Images Osama M. Abu Zaid	208-214
31. Cryptosystem Algorithm Based on Chaotic Systems for Encrypting Colored Images Osama M. Abu Zaid	215-224
32. New algorithm for automatic visualization of metro map Somayeh Sobati Moghadam	225-229
33. Optimizing the Number of Neighbors in Trust Based Recommender Systems Harmeet Kaur and Deepali Jain	230-238
34. The Difference Between Dwelling And Home In Architecture Shadab Shidfar	239-243
35. A Novel Design of Ultra-Wideband Antenna Using Slot-Based Structures Mohamed Taouzari, Ahmed Mouhsen, Jamal El Aoufi and Jamal Zbitou	244-247
36. Criticality estimation of IT business functions with the Business Continuity Testing Points method for implementing effective recovery exercises of crisis scenarios Athanasios Podaras and Tomas Zizka	248-256
37. A Metric Based Approach to Extract, Store and Deploy Software Reusable Components Effectively Muhammad Ilyas, Mubashir Abbas and Khansa Saleem	257-264
38. Sentiment Analysis of Equities using Data Mining Techniques and Visualizing the Trends Shradha Tulankar, Rahul Athale and Sandeep Bhujbal	265-269
39. A Shadow Detection and Removal from a Single Image Using LAB color Space Ashraful Huq Suny and Nasrin Hakim Mithila	270-273
40. Software Reliability Prediction using Artificial Techniques Rita G. Al Gargoor and Nada N. Saleem	274-281
41. Spectral Subtraction for Speech Enhancement in Modulation Domain Muhammad Tariq Sadiq, Noman Shabbir and Wlodek J. Kulesza	282-285
42. Spectrum Hole Prediction based on Historical Data: A Neural Network Approach Bara'u Gafai Najashi, Feng Wenjiang and Mohammed Dikko Almustapha	286-293
43. The Effect of Data Hiding at Various Bit Positions on Audio Stegnography in DCT Domain Sahib Khan, Umar Said, Ejaz Ahmad, Fazlullah and Mukhtair Ali	294-297
44. The Effect of Various Number of Least Significant Bits substitution in Audio using Discrete Cosine Transform Sahib Khan, Umar Said, Ejaz Ahmad, Fazlullah and Mukhtair Ali	298-302

45. The True Powers of Multi-core Smartphones Dhuha Basheer Abdullah and Mohammed M. Al-Hafidh	303-307
46. An Expert System for Diabetic Microvascular Complication Diagnosis Truc Ly Thi Nguyen and Nhon Van Do	308-316
47. Implementation of Information and Communication Technology in Inclusive Education in Nigeria: Future Perspective Eskay M., Ezegbe, N., Anyanwu, J. and Ikwumelu, S. N.	317-325

The Design of a Dual-band antenna for GPS/Beidou Applications

Shenzheng Zuo¹, Luyao Shi², Encheng Wang²

¹ Beijing Metstar Radar Co. Ltd
Beijing, China

² Information Engineering College, North China University of Technology,
Beijing, China

Abstract

A dual-band circularly polarized microstrip patch antenna for GPS and Beidou applications is proposed and simulated. Compared to existing navigation antennas, this antenna features dual GPS L1 and Beidou L band operation, circular polarization, and excellent near hemispherical coverage. The proposed antenna is formed by two layers of circular patches which is used M-probe as its feed network. The measured results confirm the validity of this design, which can meet the requirement of GPS and Beidou application.

Keywords: *global positioning system, Beidou, circularly polarized, dual-band, navigation antenna.*

1. Introduction

Global navigation satellite system (GNSS) refers to a kind of space navigation technology for navigation and positioning on the ground, sea and space using in the sky of the navigation satellite. With the establishment of China's Beidou satellite COMPASS networks, and with the United States GPS, Russia GLONASS and the European Union GALILEO, they are building the world four big global navigation and positioning system [1]. Also with the development of integrated navigation technology of multi-mode satellite, the design of satellite receiving antenna which can receive multiple frequency signals has received extensive attention. In view of the right-hand circular polarization navigation signal which is launched by the navigation satellite, it requires that in the working band the navigation terminal antenna has good right-hand circular polarization characteristics [2-3]. In order to meet the requirements of multi-mode and multi-band navigation technology, it needs to design an ultra broadband, multi-band and circular polarization antenna.

Microstrip antenna has low profile, light weight, low cost and easy production characteristics, but the inherent defect is narrow band. Therefore how to realize broadband of microstrip antenna with circular polarization characteristics is a hot research topic in recent years [4-5]. In the premise, this paper will study and discuss a navigation antenna

structure, focuses on the study of microstrip antenna with circular polarization and miniaturization technology, designed a kind of dual-band antenna structure model. The antenna can work at the L frequency of Beidou, and GPS L1 frequency band

In this paper, we present a novel M-probe feed stacked patch antenna to cover all two bands. It is including the design of the antenna feed network, selection of antenna's feeding methods, simulation and optimal antenna patch, at last analyzing the dielectric substrate, the feed probe, radiation effects of patch shape and height and other parameters on the antenna performance [6-7]. The antenna is using 3dB hybrid as its feeding network, and ensure that the requirements of circular polarization characteristics. It has a great application value; In addition, our proposed antenna is using air as its medium, there by achieving a reduction in design complexity, cost, and assembly time. After continuous processing, testing, and improvement, the antenna's axis ratio, pattern, gain characteristics are improving. So the antenna can meet the broadband, dual-band and circular polarization requirements [8]

2. Antenna Structure and Design

The geometry of the proposed dual-band antenna is shown in Figure 1. The antenna consists of two stacked patches, two M-probe and a 3dB hybrid. Using air as medium, greatly improving the impedance of the antenna, for the more it reduces the production cost, convenient for later processing test. The 3dB hybrid medium plate of dielectric constant for the general quantity with $\epsilon_r=3.48$. As shown in Figure 1, the feed network is composed of a directional coupler. This directional coupler will provide the required two equal signals and a 90 degrees phase difference. The 3dB hybrid is placed a circular version, its semi diameter is $R=52\text{mm}$ and thickness is $H=3\text{mm}$.

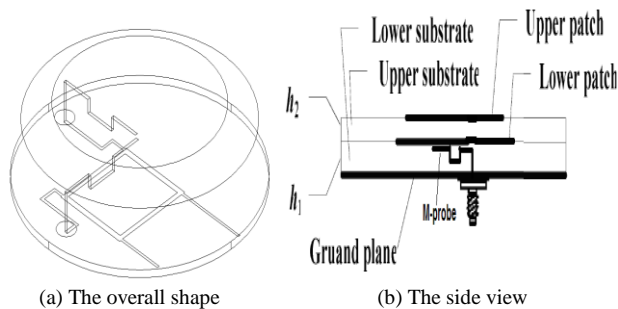


Fig. 1 proposed circular patch Antenna.

The antenna using M-probe coupling feed, due to the adoption of the structure for the antenna, it provides multi-frequency and broadband characteristics [9]. The height of this probe is $h=12\text{mm}$. The top of the M-probe are two circular radiating patches. The lower patch with radius $R1$ works at a lower frequency, while the upper patch with radius $R2$ works at a higher frequency. The height of these two layer patches are $h1$ and $h2$, respectively [10].

In order to facilitate the design procedure, first the resonant frequencies were computed using the cavity model. The final dimensions were obtained using Ansoft HFSS 10.0. The patch sizes were chosen to achieve resonance at the L band. The following antenna parameters were obtained using Ansoft HFSS 10.0: $R1=48\text{mm}$, $R2=35\text{mm}$, $h1=21\text{mm}$ and $h2=29.5\text{mm}$.

3. Experimental Result and Discussion

The return loss of the proposed antenna is shown in Figure 2. At resonant frequencies of 1.26 GHz (Beidou) and 1.575 GHz (GPS) the antenna had return loss at -14.65 dB and -24.3 dB respectively. At low frequency the simulated impedance bandwidth (10dB return loss) is 30MHz from 1.24 GHz to 1.27 GHz and at high frequency the impedance bandwidth is 50MHz from 1.54 GHz to 1.59 GHz.

As shown in Figure 3, the simulated 3 dB axial ratio at L frequency band. This kind of antenna's axial ratio has good dual-band characteristics. It is operating GPS L1 frequency and Beidou L frequency band, respectively. As shown in Fig.3, the 3dB axial ratio bandwidth of GPS L1 frequency is 90MHz from 1.5 GHz to 1.59 GHz and the Beidou L frequency 3dB axial ratio bandwidth is 60MHz from 1.21 GHz to 1.27 GHz. The simulation results of surface, this type of antenna can meet the requirements of multimode satellite navigation technology.

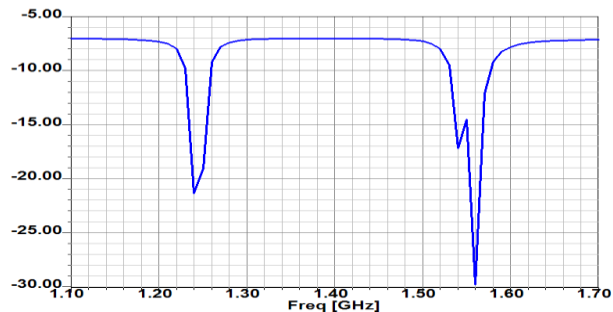


Fig. 2 Return loss of the proposed antenna.

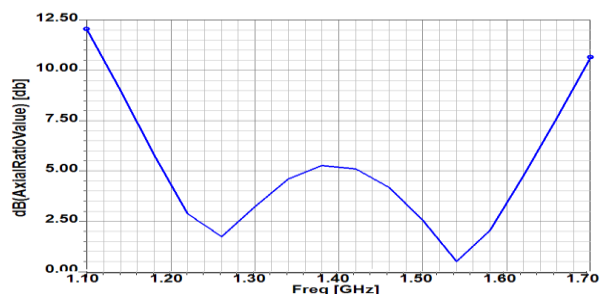


Fig. 3 Axial ratio of the antenna.

Fig 4(a/b) shows the radiation patterns of the antenna at GPS L1 frequency band. The Fig 4(a) shows the E radiation pattern plane and the Fig 4(b) show the H plane. The antenna radiation pattern has good directional diagram as shown in Fig 4. The right-hand circular polarization (RHCP) gain is 9.088 dB at 1.575 GHz. The gain difference between left-hand circular polarization (LHCP) and RHCP is greater than 13.681 dB at L1 frequency at the broadside direction. The simulation results are based on Ansoft HFSS.

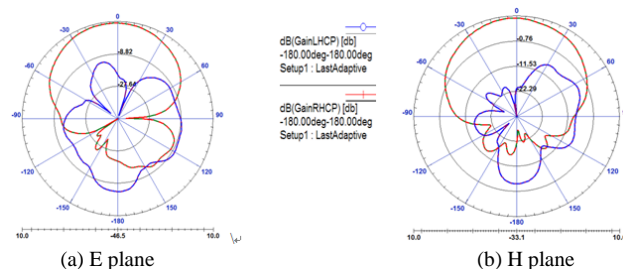


Fig. 4 Simulated Circularly Polarized Radiation Patterns at 1575MHz.

Fig 5(a/b) shows the radiation patterns of the antenna at Beidou L frequency band. The Fig 5(a) shows the E radiation pattern plane and the Fig 5(b) show the H plane. The right-hand circular polarization (RHCP) gain is 8.560 dB at 1.26 GHz. The gain difference between left-hand circular polarization (LHCP) and RHCP is greater than 19.897 dB at L frequency at the broadside direction.

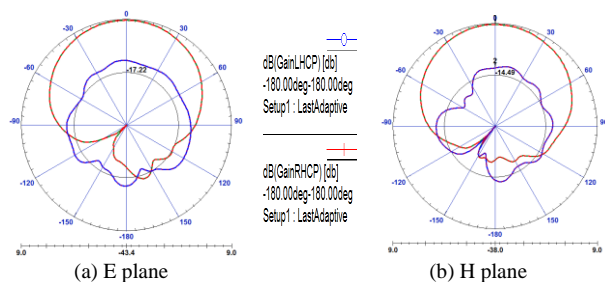


Fig. 5 Simulated Circularly Polarized Radiation Patterns at 1.26GHz.

4. The Measurement Result

According to the design and processing of a pair of a dual-band circular polarization microstrip patch antenna. The antenna is a M type probe coupled feed, using air as its medium. The proposed antenna is fabricated and measured. Fig.6 is a photograph of the assembled antenna. Fig.7 shows the measured result of the proposed antenna.



Fig. 6 Photograph of the fabricated antenna.

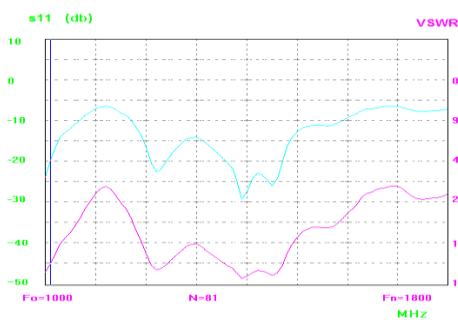


Fig. 7 Measured S11 and VSWR.

The return loss of this antenna and its VSWR (Voltage Standing Wave Ratio) are shown in Fig.7. Although good matching is observed from 1.1 to 1.7 GHz, not all of them are due to good antenna matching. Because the truncated corner is used, the 50-load can absorb all reflected power even when the antenna works at mismatched frequency points.

5. Conclusions

A novel dual-band antenna structure operated at GPS and Beidou frequencies is presented in this letter. It has impedance bandwidths larger than 50MHz and 30MHz at the GPS L1 and Beidou L frequency band, respectively. Measured gains at the broadside direction at GPS L1 and Beidou L are about 10dB and 8.87 dB, respectively. The method of varying the axial ratio is also discussed in this letter.

References

- [1] Giremus, A. Tourneret, J.-Y. Calmettes, V. "A Particle Filtering Approach for Joint Detection/Estimation of Multipath Effects on GPS Measurements", IEEE Transactions on Signal Processing, Vol.55 (4), pp.1275 – 1285, 2007.
- [2] A. Lai, T. Itoh, and C. Caloz, "Composite right/left-handed transmission line meta materials," IEEE Microwave Magazine, vol. 5, no. 3, pp. 34-50, 2004.
- [3] Zhi-Hong, Qing-Xin Chu. Circular Polarized Antenna for GPS and CNSS Applications Proceedings of Asia-Pacific Microwave Conference, 2010.
- [4] ShiChai Chen, GuangCong Liu,et.al., "Compact Dual-Band GPS Microstrip Antenna Using Multilayer LTCC Substrate", IEEE Antennas and Wireless Propagation Letters, vol.9, pp.421-423, 2010.
- [5] J.-H. Oh, Y.-P. Hong, and J.-G. Yook, "Dual circularly-polarized stacked patch antenna for GPS/SDMB," in Proc. Int. Conf. Microw. Millim. Wave Technol., 2010, pp. 252–255.
- [6] S. Liu, and Q. X. Chu, "A novel dielectrically-loaded antenna for GPS/CNSS dual-band applications," 2008 International Conference On Microwave and Millimeter Wave Technology, pp. 990-992, 2008
- [7] Ghobadi, A.; Dehmollaian, M., "A Printed Circularly Polarized Y-Shaped Monopole Antenna", IEEE Antennas and Wireless Propagation Letters, Vol.11, pp: 22-25, 2012.
- [8] The-Nan Chang; Jyun-Ming Lin; Chen, Y.G., "A Circularly Polarized Ring-Antenna Fed by a Serially Coupled Square Slot-Ring", IEEE Transactions on Antennas and Propagation, Vol.60 (2), pp. 1132-1135, 2012.
- [9] Lau, K.L.; Luk, K.M., "A Wideband Dual-Polarized L-Probe Stacked Patch Antenna Array", IEEE Antennas and Wireless Propagation Letters, Vol.6, pp.529 – 532, 2007.
- [10] Jun-Hwa Oh, Young-Pyo Hong. Dual Circularly-polarized stacked patch antenna for GSM/SDMB. Antennas and Propagation Socirty International Symposium, IEEE,2008.

Dr Shenzheng Zuo graduated from Beijing University of Posts and Telecommunications. He is an engineer of Beijing Metstar Radar Co.Ltd.

Luyao Shi is a graduate student of Information Engineering College of North China University of Technology.

Dr Encheng Wang graduated from Dalian Maritime University, He is a teacher of Information Engineering College of North China University of Technology.

Architecture Exploration of Multicore Systems-on-Chip using a TLM-based Framework

Mona Safar¹, Magdy A. El-Moursy², Mohamed Abdelsalam³, and Ashraf Salem⁴

¹ Ain Shams University, Cairo, Egypt
^{2,3,4} Mentor Graphics Corporation, Cairo, Egypt

Abstract

A framework for TLM architecture exploration of multi-core systems is presented. Starting with a Task Precedence Graph (TPG) as a design entry, different architectures with different number of processor cores, number of busses, task-to-processor and channel-to-bus mappings are automatically generated. The viability and potential of the proposed approach is demonstrated by an illustrative example.

Keywords: TLM, SoC, Architecture Exploration, Multi-Core

1. Introduction

With the increasing complexity of multi-core systems, it is necessary to have an easy, efficient, and automatic technique for fast analysis of architecture trade-offs in early stages of the design [1]. Transaction Level Modeling (TLM) provides a fast, simple and powerful methodology to enable early exploration of possible architectures [2].

Several TLM-based design space exploration methodologies for multi-core systems have been developed. MULTICUBE presents a framework for design space exploration of multi-core systems based on design parameters [3]. The multi-processor platform is simulated based on Cycle Accurate (CA) TLM. The system requires an architect engineer to explore different strategies, metrics and constraints. The research in [4] targets the architecture exploration of inter-subsystem communication in Multiprocessor Systems-on-Chip MPSoCs with A Kahn Process Network (KPN) as the design entry. A Virtual Platform (VP) level, where software is executed by Instruction Set Simulators (ISS) and hardware cores are modeled using CA models, is adopted. CA modeling slows-down simulation which increases the time-to-market.

In this paper, a new Approximately-Timed (AT) TLM-based framework for architecture exploration of specialized multi-core system architecture is presented. Starting from Task Precedence Graphs (TPG) different architectures are explored. The presented technique allows visualization and easy analysis of throughput, latency, and utilization metrics of various resources in the system.

2. TLM Multi-Core Architectural Exploration Framework

The proposed framework for TLM-architecture exploration of specialized multi-core system architecture is shown in Figure 1. The design entry is the software application specified as a set of communicating tasks exchanging data through blocking channels expressed as a Task Precedence Graph (TPG).

A scheduling algorithm [5] is applied to solve the task allocation/scheduling problem to obtain the optimum schedule on a multiprocessor system while reducing the number of processors in the target system. The algorithm also resolves conflicts in the communication channels. The algorithm provides different architecture options [modeled as Architecture Level Models (ALM)] that satisfy the optimum schedule length with different number of processor cores. Each ALM represents an architecture structure of the system which is specified by the number of cores, the number of busses, task-to-core mapping, and channel-to-bus mapping.

Once the ALM file is generated, the target multi-core system and its simulation dynamics are automatically built using scalable transaction-level modeling methodology [6]. Functionality, communication, timing, and power are all separated aspects in the presented methodology. Each processor core has Local Memory (LM) that is used for the communication among the tasks on the same core. Each CPU core is the initiator for the required read/write transactions based on the running tasks. The tasks communicate using the shared memory through specific memory addresses. All models are instantiated from a library of fast generic models.

Each CPU core is modeled using a single SC_THREAD that runs different tasks according to the provided schedule. For each task, the thread is suspended for the specified execution time based on the computation cost of the running task. State-based power [7] is used to model the computational power for tasks which are running on each processor core. A state variable is changed whenever a new scheduled task start running on the core. The power consumption is evaluated

based on the value of this state variable and is updated each time this state variable is refreshed.

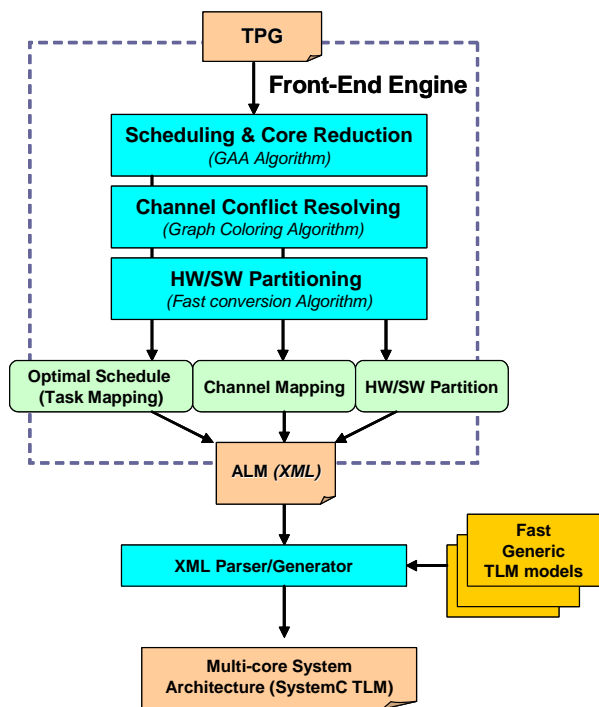


Fig. 1 TLM multi-core system architecture exploration framework.

3. Illustrative Example

A practical example is shown in Figure 2. The tasks computation cost as execution time in cycles and consumed energy in nJ are listed. The communication time for all communication channels is 200 cycles with energy consumption of 38μJ per bit. The scheduling algorithm provides two ALM models using two-core and three-core architectures with the same schedule length. A single core system was not sufficient to achieve the target latency.

The task mapping is illustrated for both architectures as shown in Figure 3. For the three-core architecture, two buses are needed to resolve the channel conflicts (Ch1 conflicts with Ch3 and Ch2 conflicts with Ch4). Using graph coloring algorithm, channels Ch1 and Ch2 are mapped to Bus1 and channels Ch3 and Ch4 are mapped to Bus2. The TLM-based architecture of the 3-core and the 2-core systems are plotted in Figure 4.

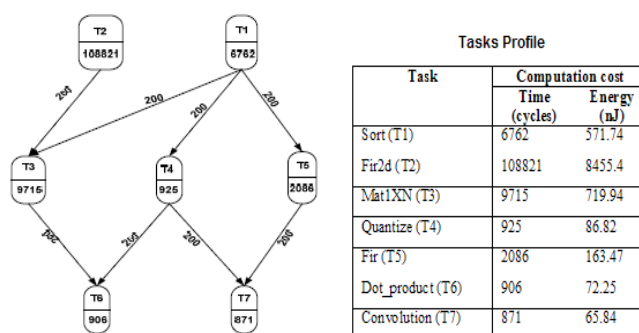


Fig. 2 TPG for a practical system.

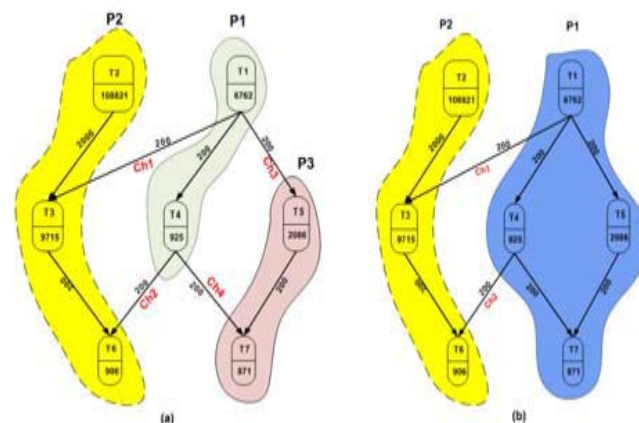


Fig. 3 Different architectures tasks mapping on (a) 3 cores, (b) 2 cores.

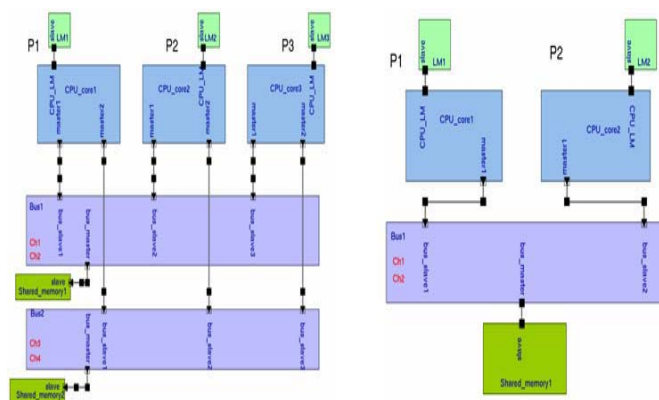


Fig. 4 TLM-based architecture. (a) 3-core, (b) 2-core.

For each architecture, simulation runs in AT mode which provides several orders of magnitude speedup over cycle accurate models while producing sufficiently accurate simulation results. For SystemC elaboration phase, LT (Loosely Timed) simulation mode, where no timing constraint is taken in account, is used for fast system bring-up. A clock speed of 100 MHz is used. The delay and power are ignored since they are negligible. The GANTT (scheduling chart) is plotted to illustrate the tasks execution schedule on the cores as

shown in Figure 5. Different statistic information for various data transfers are determined to help the designer to compare power consumption and the resource idle time of different architectures as shown in Table I.

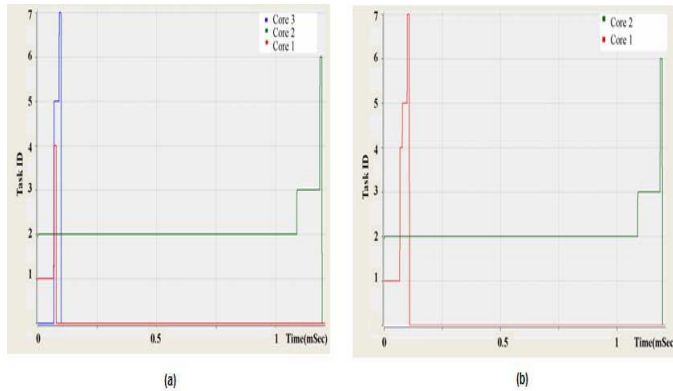


Fig. 5 GATT chart (a) 3-core, (b) 2-core.

Table 1: Comparing 2-cores and 3-cores architecture

<i>Performance criterion</i>		3-core	2-core
Power (mW)	Dynamic	1430	1059
	Clock	13	7
	Leakage	3	1
	Total power	1445	1067
Resource idle time (%)	Processor cores	63	64
	Busses	72	86

5. Conclusions

A new framework for TLM-based architecture exploration for multi-core system starting from TPG is presented. TLM provides fast and still accurate efficient exploration methodology Separating functionality, timing, and power aspects reduces the modeling effort and speeds exploring different architectures. The proposed approach not only contributes in dramatically decreasing the exploration time, but also eases design understanding, evaluation and analysis.

References.

- [1] G. Martin. "Overview of the MPSoC Design Challenge," in Proc. Of DAC'06, pp. 274-279, July 2006.
- [2] F. Ghenassia, "Transaction-Level Modeling with SystemC: TLM Concepts and Applications for Embedded Systems," Published by Springer, Netherlands, 2005.
- [3] C. Silvano, W. Fornaciari, G. Palermo, V. Zaccaria, F. Castro, M. Martinez, R. Zafalon, S. Bocchio, M. Wouters, G. Vanmeerbeeck, P. Avasare, C. Couvreur, L. Onesti, C. Kavka, A. Turco, U. Bondi, G. Mariani, E. Villar, H. Posadas, C. Y. q. Wu, F. Dongrui, Z. Hao and T. Shubin. "MULTICUBE: Multi-objective design space exploration of multi-core architectures." In IEEE Computer Society Annual Symposium on VLSI - ISVLSI, Kefalonia, Greece, July 2010.
- [4] Hao Shen; Petrot, F.; , "MPSoC Communication Architecture Exploration Using an Abstraction Refinement Method," VLSI Design, 2008. VLSID 2008. 21st International Conference on , vol., no., pp.403-408, January 2008
- [5] H. Youness, K. Sakanushi, Y. Takeuchi, A. Salem, A. Wahdan and M. Imai, "Optimal Scheme for Search State Space and Scheduling on Multiprocessor Systems", IEICE Transaction On Fundamentals of Electronics, Communications and Computer Sciences, Vol. E92-A, No.4, pp.1088-1015, April 2009.
- [6] Mentor Graphics® Vista Architect™ Tool [Online]: Available: http://www.mentor.com/products/esl/design_verification/vista_architect
- [7] R. A. Bergamaschi, Y.W. Jiang., "State systems-on-chip," Design Automation Conference, pp. 638- 641, June 2003.

Mona Safar received the B.S, M.S, and Ph.D. degrees in computer engineering from the Ain Shams University, Cairo, Egypt, in 2004, 2007 and 2011, respectively. She is currently an assistant professor in the Computer and Systems Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt. Her research interests include reconfigurable computing, application specific architectures, computer-aided design and system-level design.

Magdy A. El-Moursy was born in Cairo, Egypt in 1974. He received the B.S. degree in electronics and communications engineering (with honors) and the Master's degree in computer networks from Cairo University, Cairo, Egypt, in 1996 and 2000, respectively, and the Master's and the Ph.D. degrees in electrical engineering in the area of high-performance VLSI/IC design from University of Rochester, Rochester, NY, USA, in 2002 and 2004, respectively. In summer of 2003, he was with STMicroelectronics, Advanced System Technology, San Diego, CA, USA. Between September 2004 and September 2006 he was a Senior Design Engineer at Portland Technology Development, Intel Corporation, Hillsboro, OR, USA. During September 2006 and February 2008 he was assistant professor in the Information Engineering and Technology Department of the German University in Cairo (GUC), Cairo, Egypt. Dr. El-Moursy is currently Staff Engineer in the Mentor Graphics Corporation, Cairo, Egypt. His research interest is in Networks-on-Chip, interconnect design and related circuit level issues in high performance VLSI circuits, clock distribution network design, and low power design. He is the author of more than 40 papers, four book chapters, and two books in the fields of high speed and low power CMOS design techniques and high speed interconnect, NoC and SoC.

Mohamed AbdElSalam received his B.Sc. and M.S Degree from Ain-Shams University, Cairo, Egypt, and Doctor of Information Science and Technology from Osaka University, Osaka, Japan. He joined Mentor Graphics 1998-2002 working in SW development of circuit simulation and IC layout tools, and development of ModuleWare library in FPGA Advantage/HDS tool, and again in 2008 to present, in Global R&D Egypt MED solutions as Technical lead, working on hardware emulation targets, Memory softmodels and Transaction/Virtual based Solutions. His research interest is in HW/SW co-design, System level design, and HW modeling of communication protocols IPs.

Ashraf Salem is Engineering Director in Mentor Graphics Egypt. He manages a group of 110 engineers working in the development of Emulation, Simulation Embedded Systems and Automotive products. Dr. Salem obtained his Ph.D. from Grenoble University, France in 1992. He got his B.Sc. and M.Sc. in Computer Engineering from Ain Shams University in 1983, 1987 respectively. He was the CEO of the Technology Innovation and Entrepreneurship Center (TIEC) and professor of Computer Engineering, Faculty of Engineering, Ain Shams University. Dr. Salem participated in the establishment of ANACAD branch in Egypt in 1995 that then has been acquired by Mentor Graphics and became one of the largest multinational development centers. Dr. Salem is participated in the establishment of Software Engineering Competence Center and Information Technology Institute in Egypt. He published more than 100 scientific articles in the fields of Computer Aided design of Digital circuits. He chaired the technical committees in a number of international conference,

and he supervised more than 20 PhD and M. Sc. thesis in digital design and Embedded systems. Also, he participated in a number of international research projects and developed one of the pioneer research product for circuit verification in the eighties.

Three-dimensional Numerical Study of Natural Convection in a Cubical Enclosure with Two Heated Square Sections Submitted to Periodic Temperatures

Lahoucine Belarche¹, Btissam Abourida² and Touria Mediouni³

¹ National School of Applied Sciences, Ibnou Zohr University
Agadir, Morocco

² National School of Applied Sciences, Ibnou Zohr University
Agadir, Morocco

³ National School of Applied Sciences, Ibnou Zohr University
Agadir, Morocco

Abstract

The three-dimensional numerical study of natural convection in a cubical enclosure, filled with air was carried out in this study. Two heating square sections are placed on the vertical wall of the enclosure. The imposed heating temperatures vary sinusoidally with time, in phase and in opposition of phase. The temperature of the opposite vertical wall is maintained at a cold uniform temperature and the other walls are adiabatic. The governing equations are solved using Control volume method by SIMPLEC algorithm. The sections dimension D/L and the Rayleigh number Ra were fixed respectively at 0,2 and 10^6 . The temperature distribution, the flow pattern and the average heat transfer will be examined for a given set of the governing parameters, namely the amplitude of the variable temperatures a and their period τ_p . The obtained results show significant changes in terms of heat transfer and flow intensity, by proper choice of the heating mode and the governing parameters.

Keywords: *Three-dimensional modeling, Natural convection, Two square heated sections, variable heating, periodic temperature, dephasing.*

1. Introduction

The study of natural convection induced in cavities represents a great interest because of its direct application in various fields of engineering. Examples involving the natural convection phenomenon are building design, solar collectors, electronic and computer equipments and other applications which are referred to references [1-3]. In thermal control of electronic systems, a careful attention is necessary to ensure efficient design in order to optimize the cooling of the system. A review of the literature shows the existence of numerous numerical and analytical works on the subject. However, in most of these works, the studied configurations are two dimensional cavities, submitted to constant heating conditions. Hence, the two

dimensional natural convection in a partially heated cavity (temperature or heat flux), with one or more heating portions, has been studied previously by many authors [4-8]. However, this state is not very representative of the reality happens in most practical applications. Indeed, in many situations, the energy supplied to the system varies over time giving rise to patterns of transient or unsteady natural convection. The solar energy collectors or electronic circuits are examples of such systems. The transient natural convection is then studied previously in the case of a cavity with time variable heating (temperature or heat flux) [9-15]. The previous results show that it is impossible to predict the fluid behavior and the heat transfer induced by variable heating conditions from the results obtained with constant thermal boundary conditions. In addition, it is noted that most of the published works adopted a two dimensional model. Few works have been reported on three-dimensional natural convection in enclosures partially heated [16-20], in spite of the fact that such approach reflects the phenomenon reality and leads to important and varied results, comparing to the two dimensional cases.

Note finally that all these available works on three-dimensional natural convection induced in a cavity partially heated, considered the case of constant heating temperature.

Hence, the purpose of the present investigation is to study numerically the case of laminar natural convection in a cubical enclosure with two heating square sections placed on its vertical wall. The imposed heating temperatures are varied sinusoidally with time, either in phase or in opposition of phase. The rest of the considered wall is adiabatic while the temperature of the opposite vertical wall is maintained at a cold uniform temperature and the other walls are adiabatic. The fluid flow motion, the temperature distribution and the average heat transfer will

be examined for a given set of the governing parameters, namely the amplitude of the variable temperatures a and their period τ_p .

2. Problem formulation

A schematic of the physical problem and coordinates are shown in figure 1. It consists of a cubical enclosure with two square sections placed on its left vertical wall. The rest of the wall is thermally insulated. The heating temperatures θ_{H1} and θ_{H2} of the sections are varied sinusoidally with time as shown in figure 2, in phase and in opposition of phase. The opposite vertical wall is maintained at uniform cold temperature θ_C , and the other walls are insulated.

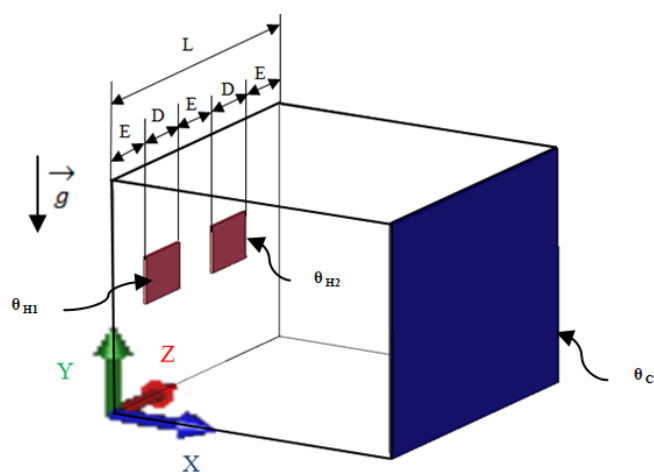


Fig. 1 Studied configuration and coordinates.

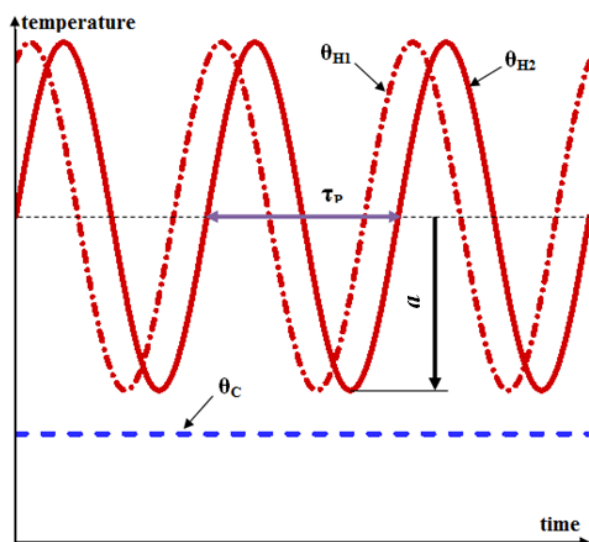


Fig. 2 Imposed thermal excitations.

The governing equations for laminar steady convection, using the Boussinesq approximation and neglecting the viscous dissipation, are expressed in the following dimensionless form

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} + \frac{\partial W}{\partial Z} = 0 \quad (1)$$

$$\frac{\partial U}{\partial \tau} + \frac{\partial}{\partial X}(UU) + \frac{\partial}{\partial Y}(VU) + \frac{\partial}{\partial Z}(WU) = -\frac{\partial P}{\partial X} + \text{Pr} \left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} + \frac{\partial^2 U}{\partial Z^2} \right) \quad (2)$$

$$\frac{\partial V}{\partial \tau} + \frac{\partial}{\partial X}(UV) + \frac{\partial}{\partial Y}(VV) + \frac{\partial}{\partial Z}(WV) = -\frac{\partial P}{\partial Y} + \text{Ra Pr } \theta + \text{Pr} \left(\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2} + \frac{\partial^2 V}{\partial Z^2} \right) \quad (3)$$

$$\frac{\partial W}{\partial \tau} + \frac{\partial}{\partial X}(UW) + \frac{\partial}{\partial Y}(VW) + \frac{\partial}{\partial Z}(WW) = -\frac{\partial P}{\partial Z} + \text{Pr} \left(\frac{\partial^2 W}{\partial X^2} + \frac{\partial^2 W}{\partial Y^2} + \frac{\partial^2 W}{\partial Z^2} \right) \quad (4)$$

$$\frac{\partial \theta}{\partial \tau} + \frac{\partial}{\partial X}(U\theta) + \frac{\partial}{\partial Y}(V\theta) + \frac{\partial}{\partial Z}(W\theta) = \left(\frac{\partial^2 \theta}{\partial X^2} + \frac{\partial^2 \theta}{\partial Y^2} + \frac{\partial^2 \theta}{\partial Z^2} \right) \quad (5)$$

The dimensionless variables used in these equations are defined as:

$$\left. \begin{aligned} (X, Y, Z) &= \left(\frac{x}{L}, \frac{y}{L}, \frac{z}{L} \right) \\ (U, V, W) &= \left(\frac{uL}{\alpha}, \frac{vL}{\alpha}, \frac{wL}{\alpha} \right) \\ \tau &= \frac{\alpha}{L^2} t \\ \theta &= \frac{T - T_r}{T_H - T_C} \quad \text{where } T_r = \frac{T_H + T_C}{2} \end{aligned} \right\} \quad (6)$$

In the above equations, the parameters Pr and Ra denote the Prandtl number, and the Rayleigh number, respectively. These parameters are defined as where

$$\text{Pr} = \frac{\nu}{\alpha} \quad \text{and} \quad \text{Ra} = \frac{g\beta \Delta T L^3}{\alpha \nu} \quad (7)$$

The hydrodynamic boundary conditions are such as the velocity components are zero on the rigid walls of the enclosure ($U = V = W = 0$). The dimensionless thermal boundary conditions associated to the governing equations are:

- Heating section 1 :

$$\theta_{H1} = 0.5 + a \sin \left(\frac{2\pi\tau}{\tau_p} \right) \quad \text{at} \quad \frac{(1-2D/L)}{3} < Z < \frac{(1+D/L)}{3},$$

$$\frac{(1-D/L)}{2} < Y < \frac{(1+D/L)}{2} \quad \text{and} \quad X = 0 \quad (8a)$$

- Heating section 2 :

$$\theta_{H2} = 0.5 + a \sin \left(\frac{2\pi\tau}{\tau_p} + \varphi_p \right) \quad \text{at} \quad \frac{(2-D/L)}{3} < Z < \frac{(2+2D/L)}{3},$$

$$\frac{(1-D/L)}{2} < Y < \frac{(1+D/L)}{2} \quad \text{and} \quad X = 0 \quad (8b)$$

- Elsewhere on the left vertical wall:

$$\frac{\partial \theta}{\partial n} = 0 \quad (8c)$$

- Right vertical wall:

$$\theta_C = -0.5 \text{ at } X = 1 \quad (8d)$$

- Other vertical and horizontal walls

$$\frac{\partial \theta}{\partial n} = 0 \quad (8e)$$

Where n is the normal direction to the considered wall and a , τ_p and φ_p are respectively the amplitude of the heating temperature, its period and the dephasing between the two heating temperatures. The amplitude a and the period τ_p is defined respectively, as:

$$\tau_p = \frac{\alpha}{L^2} t_p \text{ and } a = \frac{A}{T_H - T_C} \quad (9)$$

The average heat flux on the cold wall is calculated at each time step by:

$$Nu = \left(\int_0^1 \int_0^1 \frac{\partial \theta}{\partial X} \Big|_{x=1} dYdZ \right) \quad (10)$$

The mean Nusselt number characterizing the heat transfer through the right wall is evaluated as:

$$\overline{Nu}_C = \frac{1}{\tau_{Nu}} \int_0^{\tau_{Nu}} \left(\int_0^1 \int_0^1 \frac{\partial \theta}{\partial X} \Big|_{x=1} dYdZ \right) d\tau \quad (11)$$

τ_{Nu} represent the period of Nusselt temporal variations.

3. Numerical method

The Navier-Stokes and energy equations are discretized by the finite volume method developed by Patankar [21] adopting the power law scheme. To overcome the difficulty associated with the determination of the pressure, we suggest solving the equations of conservation of momentum coupled with the continuity equation using the SIMPLEC algorithm. For solving the algebraic system obtained after discretization of partial differential equations, the Alternating Direction Implicit scheme (ADI) is used. Tridiagonal system obtained in each direction is solved using the THOMAS algorithm. Convergence of the numerical code is established at each time step according to the following criterion, which fixes the relative difference between the field variables ϕ ($= U, V, W, T, P$),

in successive time steps (n and $n+1$) less than 10^{-5} .

$$\sum_{i,j,k=1}^{i_{\max}, j_{\max}, k_{\max}} \frac{|\phi_{i,j,k}^{n+1} - \phi_{i,j,k}^n|}{|\phi_{i,j,k}^n|} \leq 10^{-5} \quad (12)$$

i, j and k are the grid positions.

The flow pattern obtained in the case of constant heating was used as the as initial condition for the numerical calculations conducted in the case of variable heating temperature in order to favour the convergence of the numerical code.

To check the effect of grid size, preliminary tests were conducted for different combinations of the governing parameters. Table 1 shows the values of time averaged Nusselt number, Nu_c obtained for different grid sizes in the case $D/L = 0.2$, $\tau_p = 0.8$, $a = 0.8$ and $Ra = 10^6$.

To obtain a fine mesh at the active walls ($X = 0$ and $X = 1$), the computational domain was discretized by adopting a non-uniform mesh in the X direction and uniform in directions Y and Z .

The grid size effect on Nusselt number calculated using Eq.10 is presented in table 1 in the case of constant heating temperatures. The results induced by the grid $41 \times 41 \times 41$ differs by less than 2% from those obtained with a refined grid $71 \times 71 \times 71$. Hence, the non uniform staggered grid of $41 \times 41 \times 41$ nodes was estimated to be appropriate for the present study since it permits a good compromise between the computational cost (a significant reduction of the execution time) and the accuracy of the obtained results.

Finally, the accuracy of the developed numerical code was checked by comparing the present results, obtained for constant heating temperature with those previously published by Fusegi et al. [22] in the case of cubical enclosure with a completely heated vertical wall and by Frederick et al. [19] in the case of cubical enclosure with a partially heated wall ($s/L = 0.3$). A comparison of the averaged Nusselt number Nu , and maximum values velocities U and V , in the mid-plane $Z = 0.5$ is given in Table 2 for $Ra = 10^6$. The obtained results show excellent agreement with the two references, with maximum differences not exceeding 3.38% for Nu , 2.45% and 3.87% respectively for U_{\max} and V_{\max} comparing to Fusegi's results [22] and 2.38% for Nu , 2.045% and 3, 34% respectively for U_{\max} and V_{\max} comparing to Frederick result's [19].

Table 1: Effect of the grid size on Nu_c for $D/L = 0.2$, $\tau_p = 0.8$, $a = 0.8$ and $Ra = 10^6$

Maillage	\overline{Nu}_C
31x31x31	2.06668
41x41x41	2.17545
51x51x51	2.19938
61x61x61	2.20852
71x71x71	2.21896

Table 2: Validation of the numerical code against published results in terms of Nu, U_{max} and V_{max} for $Ra = 10^6$

	Nu	U_{max}	V_{max}
Fusegi et al. [22]	8.77	0.08416	0.2922
Present work	9.066	0.08622	0.3035
Difference (%)	3.38	2.45	3.87
Frederic et al. [19]	3.4857	58.3830	151.693
Present work	3.5687	58.5024	152.1997
Difference (%)	2.38	2.045	3, 34

4. Results and discussion

In this section, we present the effects of the period, amplitude and heating mode on temporal evolutions of maximum velocity in the mid-plane of the cavity (U_{max}) and the heat loss through its cold wall (Nu_c). The average values are compared for different sets of the amplitude and the period. We also present streamlines and isotherms produced during a flow cycle for each type of heating mode in order to illustrate the behavior of the fluid motion and the heat exchange in the cavity due to the thermal excitation. Due to the big number of governing parameters, we choose to fix the Rayleigh number Ra and the heating sections dimension to respective values of 10^6 and 0.2. This choice, which characterizes adequately the fluid motion and heat transfer phenomenon in the cavity, was made after many numerical simulations conducted for different Ra and different D/L . Hence, the results presented in this paper are obtained for fixed $Ra = 10^6$, $D/L = 0.2$ and $Pr = 0.71$ (air) and varied $0 \leq a \leq 1$, $0 \leq \tau_p \leq 1$ and $\rho = 0$ or π .

4.1 Influence of the period

To illustrate the effect of the parameter τ_p , we present the temporal evolution of U_{max} (calculated in the plane $Z = 0.5$) and Nu_c for different modes of heating, $a = 0.8$ and $Ra = 10^6$. Thus, figures 3a and 3b show respectively the temporal evolutions of U_{max} and Nu_c obtained in the case of the two imposed temperatures (θ_{H1} and θ_{H2}) varying in phase for $0.1 \leq \tau_p \leq 0.4$. These figures show that all the obtained solutions are periodic, with resulting periods equal to those imposed on the exciting temperatures. We also note that the sinusoidal nature of oscillations is maintained in the Nu_c evolution, but disappears in the case of U_{max} when the period increases. The observed rate is due to the fact that the imposed temperatures have low values in a part of the cycle and gives rise to a special fluid motion as seen in the streamlines presentation (figures 6a-6b and 6g-6h). In addition, the oscillation amplitudes of U_{max} undergo significant decreases when τ_p is increased, while those of Nu_c increased with τ_p . For the same conditions as before, we examine the effect of the period in the case of two imposed temperatures (θ_{H1} and θ_{H2}) varying in opposite of phase. Thus, Figures 4a and 4b show

respectively the temporal evolutions of U_{max} and Nu_c obtained in the case of the two imposed temperatures (θ_{H1} and θ_{H2}) varying in opposition of phase for $0.1 \leq \tau_p \leq 0.4$. These figures show that the obtained solutions are also periodic. But, the resulting periods are equal to half of those imposed on the exciting temperatures. We also note the sinusoidal nature of oscillations is conserved in the evolution of Nu_c , but not for U_{max} . The maximum amplitudes of oscillation of U_{max} undergo significant decreases when τ_p is increased, while those of Nu_c increase with τ_p . These trends are similar to those encountered in the case of the two imposed temperatures varying in phase (figures 3a and 3b). However, compared to the constant heating values (dashed lines), the resulting Nu_c values seems to be greater all the time, in the case of opposition of phase variation. In addition, the velocity U_{max} shows a double periodicity which is not observed in the previous case (phase variation) and can denote a special flow motion inside cavity as it can be seen in the following paragraph.

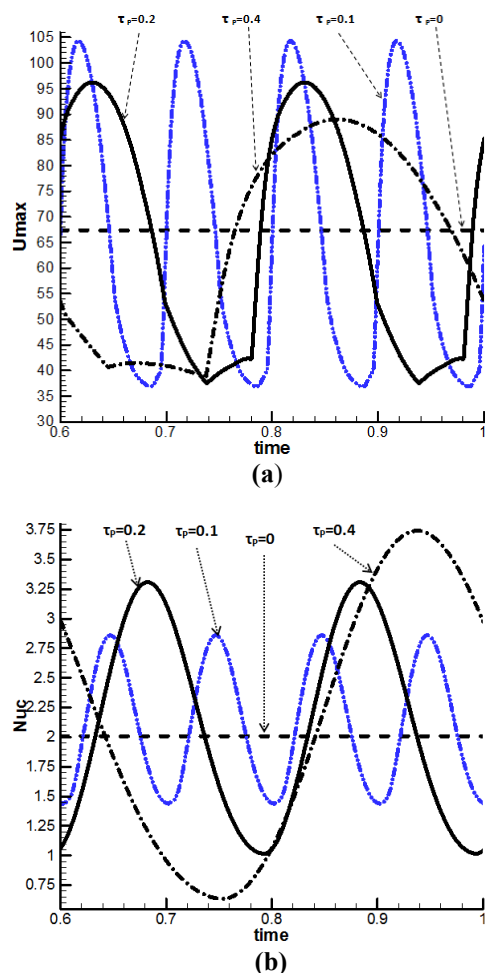


Fig. 3 Effect of τ_p for $a = 0.8$ and $Ra = 10^6$ in the case of two temperatures varying in phase: a) U_{max} and b) Nu_c .

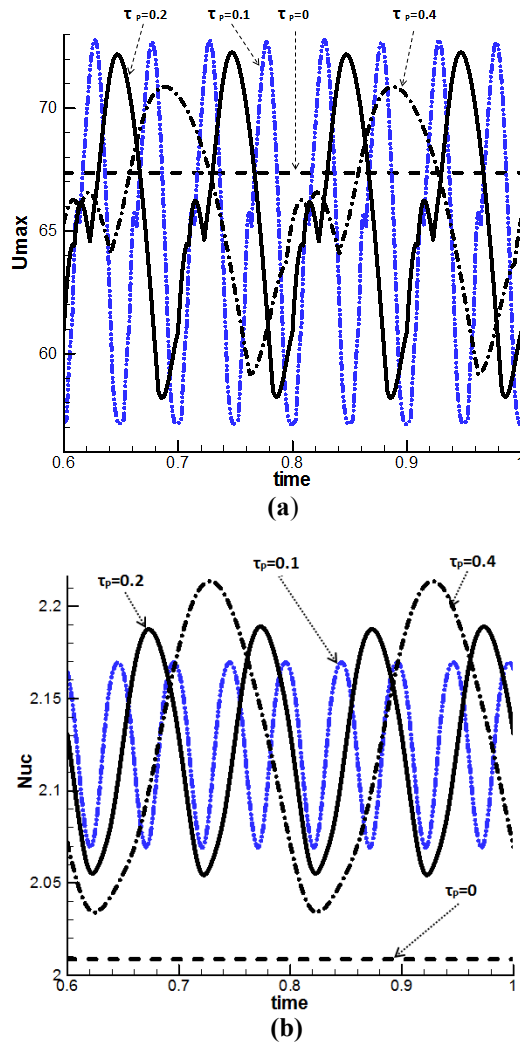


Fig. 4 Effect of τ_p for $a = 0.8$ and $Ra = 10^6$ in the case of two temperatures varying in opposition of phase: a) U_{max} and b) Nuc

4.2 Isotherms and streamlines

In order to study the details of the flow and heat transfer within the cavity, isotherms and flow lines are produced for each heating mode, during a flow cycle for $Ra = 10^6$, $D/L = 0.2$, $a = 0.8$ and $\tau_p = 0.2$. The corresponding instants are chosen during a flow cycle of velocity U_{max} (figures 5a and 5b). The heating temperatures evolution is also presented in these figures to illustrate their direct effect on the flow structure and the heat exchange.

Hence, figures 6a to 6h show isotherms (on the left) and streamlines (on the right), corresponding to the instants (a) to (h) in Figure 5a when the two imposed temperatures are varying in phase. For more visibility of the fluid motion, streamlines are presented at the mid-plane of the cavity ($X = 0.5$). Further flow presentations were produced for

different plans in the cavity, and show that the flow consists of a single cell, turning clockwise and presenting a symmetry relative to the center of the cavity ($Z = 0.5$) all the time. The cell intensity varies during the cycle.

Initially, when the imposed temperatures take their lowest values, the streamlines are straight and vertical while the heat exchange is very weak through the cavity (Figure 6a). Evolving in time, two small cells-rotating in opposite directions-appear on the upper corner of the cavity (Figure 6b). The intensity of these cells increases (figure 6c) and reaches a maximum value as it is seen in figure 6d. This instant correspond to a maximum heating period as shown in figure 5a. After that, a heating temperatures decrease leads to a flow intensity and heat exchange diminution (Figures 6f to 6h). Generally, we can note that the isotherms and the streamlines present a symmetry relative to the plan $Z=0.5$ during the entire cycle flow, when the two heating temperatures are varied in phase.

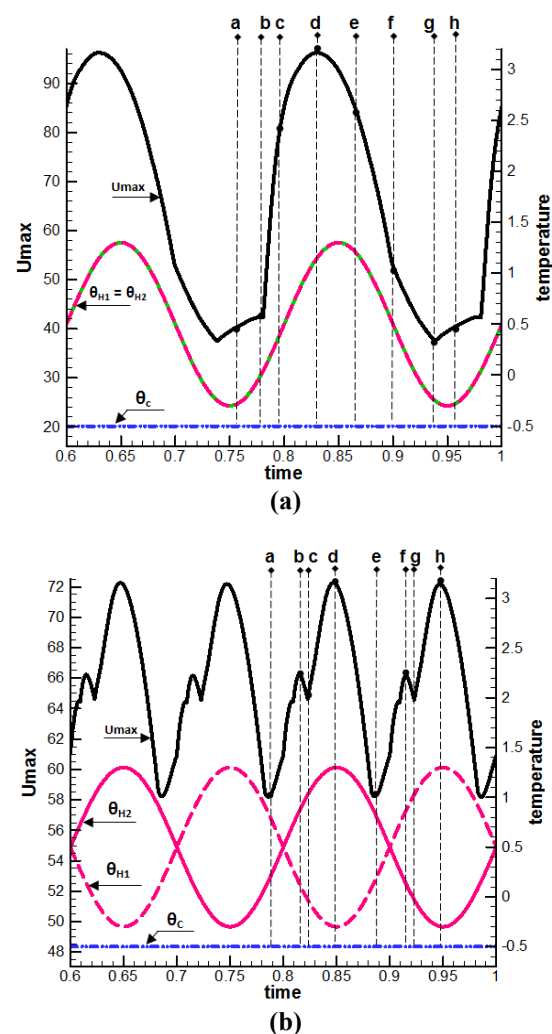


Fig. 5 Temporal variations of U_{max} for $\tau_p = 0.2$ and $a = 0.8$, and excitation temperatures: a) in phase, b) in opposition of phase.

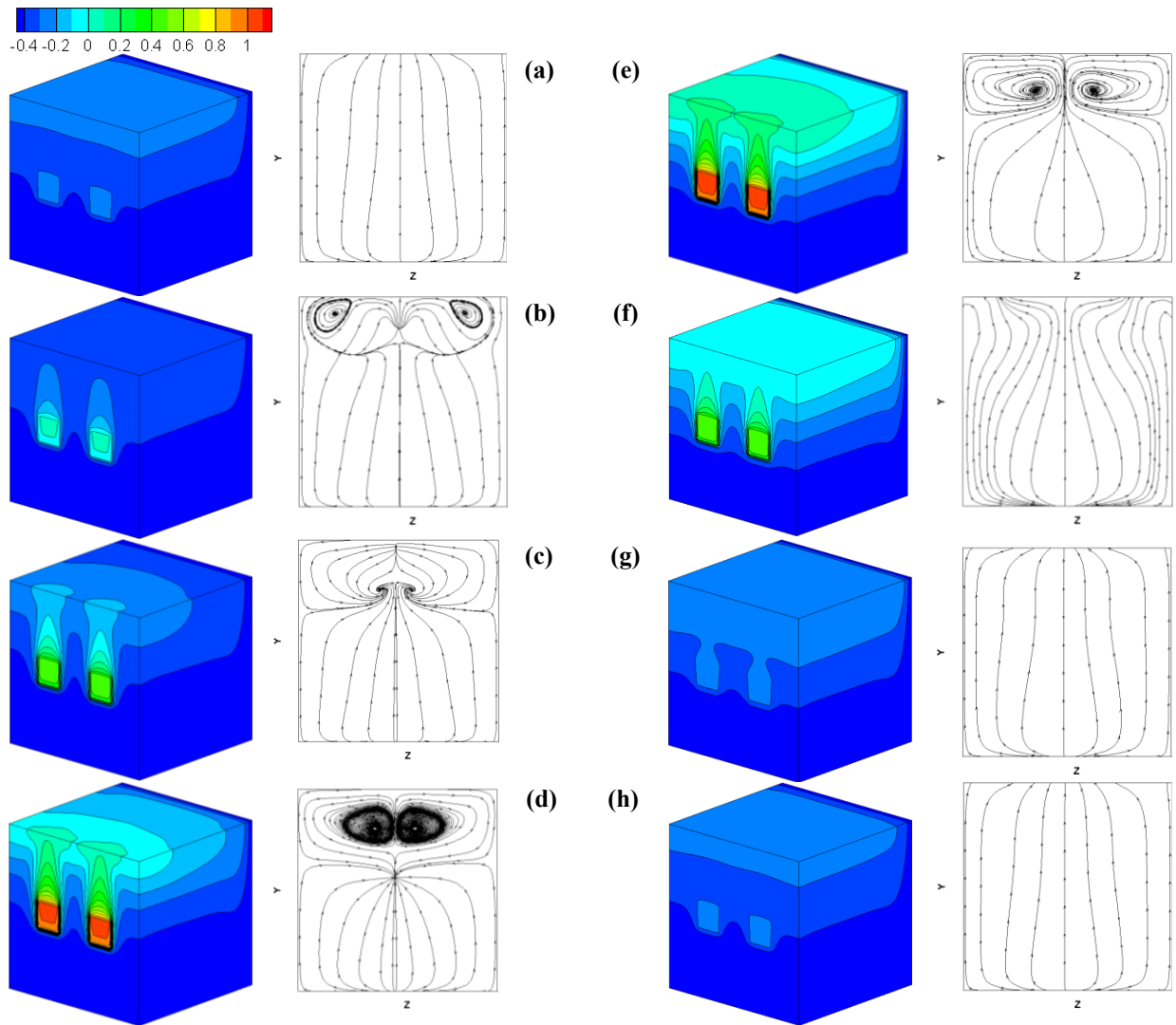


Fig. 6 Isotherms and streamlines (at plan $X = 0.5$) over one cycle for $Ra = 10^6$, $a = 0.8$ and $\tau_p = 0.2$ in the case of two temperatures varying in phase.

Figures 7a-7h show the dynamic and thermal behaviors obtained when the two imposed temperatures are varied in opposition of phase. The corresponding instants are checked out in figure 5b. It can be seen from this figure that a single heating period leads to a double resulting period of U_{max} . This behavior wasn't encountered in the case of two heating temperatures varying in phase and can explain the absence of symmetry relative to the center of the cavity ($Z = 0.5$), observed in the previous case.

Hence, the heating section becomes in turn the major source of heating and leads to recirculation cells in the top of the cavity. These cells are placed in the opposite upper corner of the cavity, according to the dominant heater section position (figures 7a-7h).

4.3 Mean values

In order to study the effect of the heating mode on mean heat transfer, figures 8a and 8b illustrate the average values of $\overline{Nu_C}$ versus the period τ_p for different amplitudes a , when the imposed temperatures are respectively varied in phase and in opposite of phase. The value corresponding to the case of constant heating source temperature ($a = 0$) are also presented in the same figures as references. It should be noted that for all the considered values of a , the presented function remain close to the permanent value in the case of weak periods. This indicates an insensitivity of the system with respect to the imposed thermal excitation.

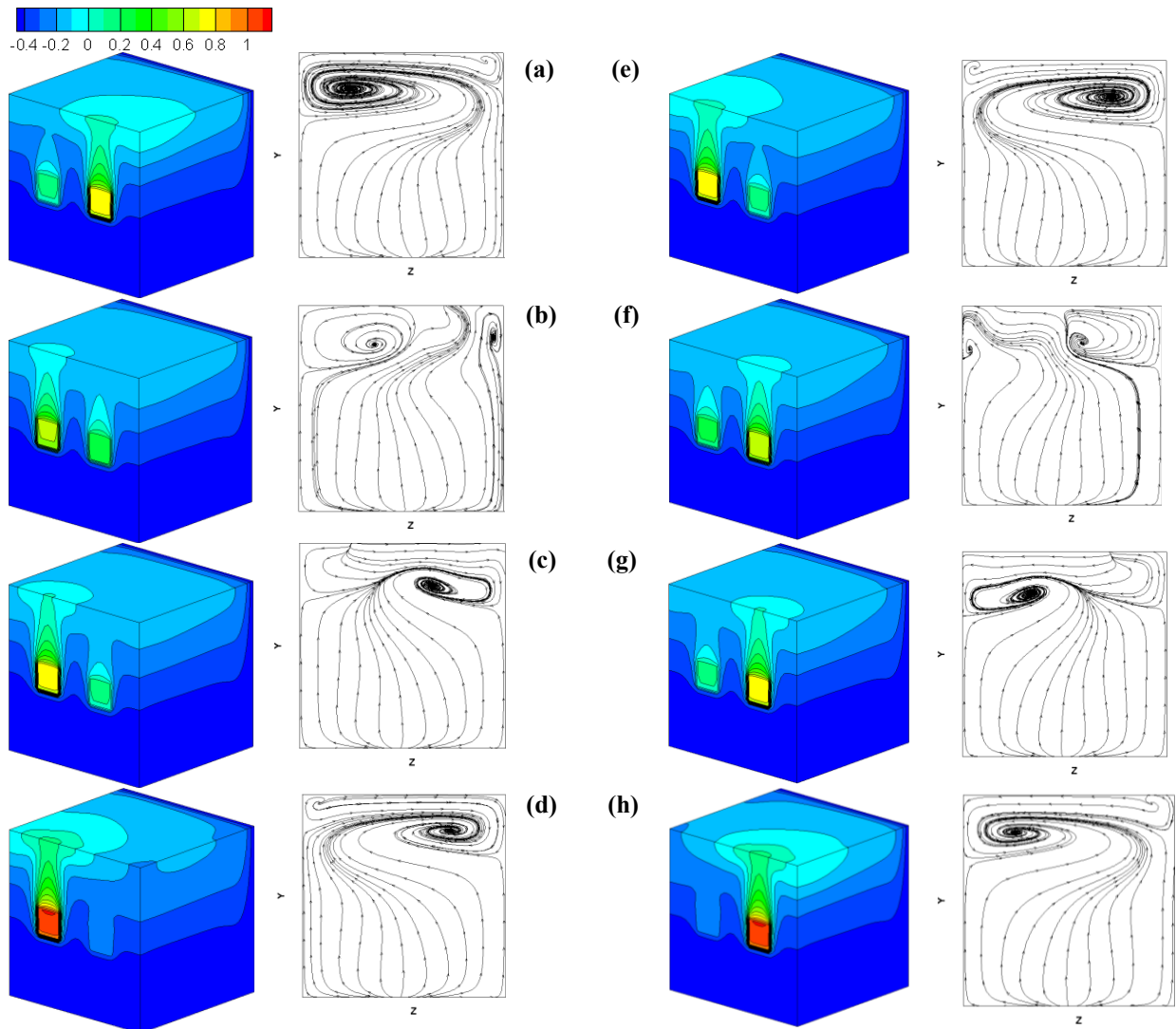


Figure 7. Isotherms and streamlines at plan $X = 0.5$ over one cycle for $Ra = 10^6$, $a = 0.8$ and $\tau_p = 0.2$ in the case of two temperatures varying in opposition of phase.

In the case of the two temperatures varying in phase (figure 8a), \overline{Nu}_c present peaks for the same value of the period of resonance ($\tau_p = 0.1$), which is not affected by the increase of the amplitude. The improvement of heat transfer with respect to the steady state value ($\overline{Nu}_c = 2.0057$) is found to increase with a and reaches 2.06%, 6.7% and 13.8% respectively for $a = 0.4, 0.8$ and 1. Above this threshold, the increase of τ_p leads to a continuous weakening of the heat transfer until reaching minimum values which is higher than the steady value. It is to be mentioned that any increase of the amplitude contributes to the enhancement of the heat transfer through the cold wall of the cavity.

When the two temperatures are varied in opposition of phase (figure 8b), variations of \overline{Nu}_c with τ_p are generally similar to those observed in the previous case. However, \overline{Nu}_c present peaks for different values of the period which increases with the amplitude. Hence, the period of resonance was found to be equal to 0.1, 0.2 and 0.4 respectively for a equal to 0.4, 0.8 and 1. The enhancement of \overline{Nu}_c with respect to the steady state value ($\overline{Nu}_c = 2.0057$) is then equal to 2.7%, 9.69% and 14.59% respectively when a is equal to 0.4, 0.8 and 1. Hence, this last mode present better heat transfer compared to the case of the two heating temperatures varying in phase.

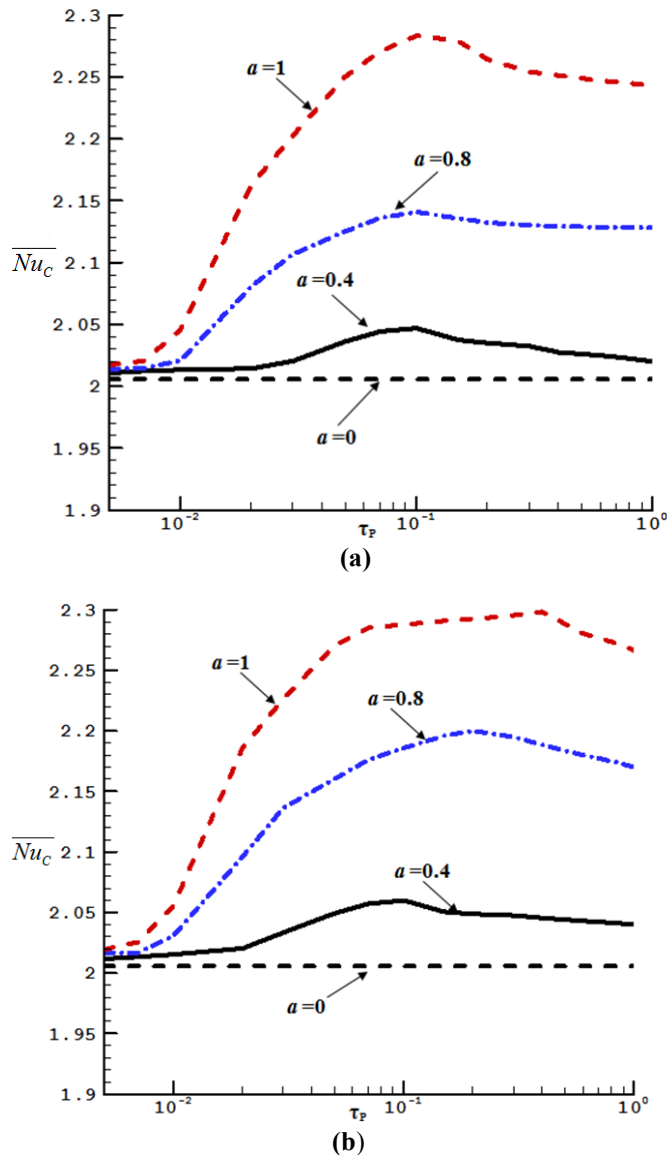


Fig. 8 Effect of τ_p for $Ra = 10^6$ and various a : a) case of the two temperatures varying in phase b) case of the two temperatures varying in opposite phase.

4. Conclusion

The numerical study of natural convection in a cubical enclosure with two heating square sections submitted to periodic temperatures has identified the following conclusions:

- The average heat transfer through the cold wall of the cavity in the case of two variable temperatures is greater than its permanent value, if we except the case of short periods and small amplitudes;
- Varying the two temperatures in opposition of phase gives rise to a better heat transfer and an important fluid

motion compared to the case of the two temperatures varying in phase.

- The resonant period, corresponding to the maximum mean heat transfer, depend on the imposed heating mode: In case the two temperatures varying in phase, it is close to the value 0.1 and independent of the excitation temperatures amplitudes, while it varies with the amplitude in the case of the two temperatures varying in opposition of phase.

Nomenclature

- A = Dimensional amplitude of the heating temperatures
- a = Dimensionless amplitude of the heating temperatures, Eq. (9)
- D = Side of the square hot sector
- g = Gravitational acceleration
- L = cavity side
- Nu = Average Nusselt number, Eq. (10)
- Pr = Prandtl number, Eq. (7)
- P = Non-dimensional pression
- p = Pression N/m^2
- Ra = Rayleigh number, Eq. (7)
- T = Température
- t_p = dimensional period
- ΔT = Overall temperature difference, $(\bar{T}_H - T_C)$
- u,v,w = dimensional velocities
- U,V,W = dimensionless velocities, Eq. (6)
- x, y, z = dimensional coordinates
- X,Y,Z = dimensionless Cartesian coordinates, Eq. (6)

Greek symbols

- α = thermal diffusivity
- θ = non-dimensional temperature, Eq. (6)
- β = volumetric thermal expansion coefficient
- τ = non-dimensional time, Eq. (6)
- τ_p = non-dimensional period, Eq. (9)
- ν = kinematic viscosity
- ρ = density
- φ_p = dephasing, Eq. (8b)

Subscripts:

- C = cold
- H = hot
- Nu = Nusselt number
- H1 = heating section 1
- H2 = heating section 2
- r = Resonance

Superscripts

- = Time averaged quantity

References

- [1] S. Ostrach, "Natural convection in enclosures", *Int. J. Heat Transfer*, Vol. 110, 1988, pp. 1175–1190.
- [2] A. Bejan, and A. D. Kraus, *Heat Transfer Handbook*. Wiley, New York, 2003.
- [3] R.J. Goldstein, et al., "A review of 2003 literature", *Int. J. of Heat and Mass Transfer*, vol. 49, 2006, pp 451-534.
- [4] H. H. S. Chu, S. W. Churchill, and C.V.S. Patterson, "The effect of heater size, location, aspect ratio, and boundary conditions on two-dimensional, laminar, natural convection in rectangular channels", *Int. J. Heat Transfer*, vol. 98, 1976, pp. 195-201.
- [5] B.L. Turner and R.D. Flack, "The experimental measurement of natural convective heat transfer in rectangular enclosures with concentrated energy sources", *Int. J. Heat Transfer*, vol. 102, 1980, pp. 237-241.
- [6] N. Nithyadevi, P. Kandaswamy and J. Lee, "Natural convection in a rectangular cavity with partially active side walls", *Int. J. of Heat Mass Transfer*, vol. 50, 2007, pp. 4688–4697.
- [7] D. Qi-Hong, "Fluid flow and heat transfer characteristics of natural convection in square cavities due to discrete source-sink pairs", *Int. J. of Heat Mass Transfer*, vol 51, 2008, pp. 5949–5957.
- [8] T. Dias and L.F. Milanez, "Optimal location of heat sources on a vertical wall with natural convection through genetic algorithms", *Int. J. Heat Mass Transfer*, vol. 49, 2006, pp. 2090–2096.
- [9] P. Vasseur and L. Robillard, "Natural Convection in a Rectangular Cavity With Wall Temperature Decreasing at a Uniform Rate", *Waerme- Stoffuebertrag*, vol. 16, 1982, pp. 199–207.
- [10] J.I. Lage, A. Bejan, "The resonance of natural convection in an enclosure heated periodically from the side", *Int J Heat Mass Transfer*, vol. 36, 1993, pp. 2027–2038.
- [11] B. Abourida, M. Hasnaoui and S. Douamna, "Convection Naturelle dans une Cavit  Carr e avec des Parois Verticales Soumises   des Temp ratures P riodiques", *Rev. Gen.Therm*, vol. 37, 1998, pp 788–800.
- [12] B. Abourida, M. Hasnaoui and S. Douamna, "Transient natural convection in a square enclosure with horizontal walls submitted to periodic temperatures", *Numerical Heat Transfer, PartA*, vol. 36, 1999, pp. 737–750.
- [13] I.E. Sarris, I. Lekakis, and N.S. Vlachos, "Natural convection in a 2D enclosure with sinusoidal upper wall temperature", *Numer Heat Transfer*, vol 36, 2002, pp. 513–520.
- [14] C.Y. Soong, P.Y. Tzeng and C.D. Hsieh, "Numerical study of bottom-wall temperature modulation effects on thermal instability and oscillatory cellular convection in a rectangular enclosure", *Int. J. of Heat and Mass Transfer*, vol 44, 2001, pp. 3855-3868.
- [15] N. Ben Cheikh, B. Ben Beya and T. Lili, "Aspect ratio effect on natural convection flow in a cavity submitted to a periodical temperature boundary", *Int. J. Heat Transfer*, vol. 129, 2007, pp 1060–1068.
- [16] I. Sezai and A. Mohamad, "Natural convection from a discrete heat source on the bottom of a horizontal enclosure", *Int. J. of Heat and Mass Transfer*, vol. 43, 2000, pp 2257-2266.
- [17] N. Ben-Cheikh, A. Campo, N. Ouertatani and T. Lili, "Three-dimensional study of heat and fluid flow of air and dielectric liquids filling containers partially heated from below and entirely cooled from above", *Int. J. Heat Transfer*, vol. 37, 2010, pp 449–456.
- [18] P.H. Oosthuizen and J. T. Paul, "Natural convection in a rectangular enclosure with two heated sections on the lower surface", *Int. J. Heat and Fluid Flow*, vol. 26, 2005, pp 587–596.
- [19] R.L. Frederick and F. Quiroz, "On the transition from conduction to convection regime in a cubical enclosure with a partially heated wall", *Int. J. Heat Transfer*, vol. 44, 2001, pp 1699–1709.
- [20] G. Wang, M. Zeng, Y.C. Ren, H. Ozoe and Q.W. Wang, "Transient heat flux measurement of natural convection in an inclined enclosure with time-periodically-varying wall temperature", *Experimental Thermal and Fluid Science*, vol. 35, 2011, pp 105–111.
- [21] S.V. Patankar, "Numerical Heat Transfer and Fluid Flow". McGraw-Hill, New York, 1980.
- [22] T. Fusegi, J.M. Hyun, K. Kuwahara and B. Farouk, "A numerical study of three-dimensional natural-convection in a differentially heated cubical enclosure", *Int. J. Heat Mass Transfer*, vol. 34, 1991, pp 1543-1557.

Lahoucine Belarche was born in Agadir, Morocco. He received a License degree from ENSET School, Mohamed 5 Souissi University Rabat, and a master degree from ENSA School, Ibn Zohr University Agadir in 2010. He is currently working toward the PhD degree in Computational fluid dynamics and thermal simulation engineering research at ENSA School, Ibn Zohr University Agadir, Morocco.

Btissam Abourida is currently a Professor (Professeur Habilit ) at the Energy and environment engineering process department of National School of Applied sciences (ENSA), Ibn Zohr University, Agadir, Morocco, since 2008. She received her PhD degree in 1999 at Cadi Ayyad University (Morocco), in Energy specialty. Her current research interests heat transfer phenomenon and renewable energy. She is responsible of energy and transfer phenomenon research team at ENSA.

Touria Mediouni is currently a Professor at the Energy and environment engineering process department of National School of Applied sciences (ENSA), Ibn Zohr University, Agadir, Morocco, since 1998. She received her PhD degree in 1994 at Cadi Ayyad University (Morocco), in Energy specialty. Her current research interests evaporation and condensation, cooling machine, solar energy. She is director of the mechanics, energy and environment laboratory LMPEE.

Characteristics of Global Calling in VoIP services:

A logistic regression analysis

Thanyalak Mueangkhot¹, Julian Ming-Sung Cheng² and Chaknarin Kongcharoen³

¹ Department of Business Administration, National Central University,
Jhongli City, Taoyuan Country, 32001, Taiwan (R.O.C.)

² Department of Business Administration, National Central University,
Jhongli City, Taoyuan Country, 32001, Taiwan (R.O.C.)

³ Department of Computer Science & Information Engineering, National Central University,
Jhongli City, Taoyuan Country, 32001, Taiwan (R.O.C.)

Abstract

We investigate the characteristics of Voice over Internet Protocol (VoIP) services in Taiwan that influence Global Calling choice, by comparing Bayesian and Maximum likelihood estimates. The results will assist operators in setting their services. The sample data are 573 sales transactions using prepaid cards issued by Taiwanese e-retailers. The dataset contains individuals who placed recommendations in 2011-2013, which are representative of the actual market. Global Calling is defined as a critical characteristic selected by consumers. Data are analyzed by logistic regression of the two approaches. The study shows that the odds ratio of Global Calling is suitably modeled by the Bayesian approach. Collectively, the results show that the key characteristic of VoIP service is the price per month and Taiwanese landline calling. Taiwanese landline calling exerts the greater positive effect on Global Calling choice. The study then discusses the implications of VoIP operator selection, focusing on Taiwan's local call market.

Keywords: *Characteristic of VoIP, Logistic regression, Bayesian estimation, Maximum Likelihood estimation, Global Calling.*

1. Introduction

International calling (hereafter called Global Calling) is offered by Internet Network Operators, including the dominant international player Skype and others such as Phone Power, ITP, Lingo, Tokbox and iCall. Consumers can make international calls at zero or very low rates by Voice over Internet Protocol (VoIP). By the end of 2011, more than 2.3 billion people worldwide could access the internet [1]. To satisfy consumers, operators attempt to improve their own infrastructure technology and provide a range of services. The majority of VoIP call services

utilize existing long distance calling, implementing voice technologies such as Global Calling.

Therefore, Global Calling is a critical characteristic by which operators can develop and increase revenue. Operators are interested in the service characteristics (attributes) that affect Global Calling services, and in which methods are applicable to small data sources. Using VoIP services, operators may strategically design and use instruments that promote Internet communication services.

In this paper, we investigate the impact of VoIP service characteristics on consumer decision in Taiwan's local call market. The accuracy of Bayesian predictors is compared that of Maximum likelihood (ML) estimates. The methodology uses the odds ratio of Global Calling. The estimated parameters provide insight into the characteristics of VoIP, and can thereby be used to improve international call performance.

The remainder of this study is organized as follows. The next section briefly reviews existing knowledge of VoIP services and the application of logistic regression to the service sector. Section 3 describes the data sources and evaluates the relationship between VoIP service characteristics and Global Calling choice by logistic regression. The results of the analysis are presented in Section 4, while Section 5 provides a summary and conclusions.

2. Literature review

The relevant literature can be broadly divided into management of the VoIP characteristics and application of

the logistic model to the service sector. We discuss each class in turn below.

2.1 VoIP characteristics

Largely as a result of globalization, electronic communication technologies have become faster, cheaper, more reliable and more effective in recent decades, and have supplanted traditional means of sending information across the globe, which are slow and often unpredictable. VoIP has become an effective means of international calling at comparatively cheap rates. This innovation in the communication industry has led to new voice technologies, such as prepaid card calling. Prepaid phone cards are compatible with any phone, anywhere at any time, and have been adapted to VoIP services, where they are used to make international calls at reduced rates. These cards have led to rapid sales increases in commercial local markets such as voipindia.co.in, aglow.sg, indosat.com, skype.com, and skype.pchome.com.tw.

This study concentrates on the role of characteristic services in telecommunication sectors, which are perceived in different ways by consumers. Previously, the relationships between VoIP choice and telephone service features have been investigated by part-worth function models, which were validated on consumer data [2, 3].

In this study, VoIP characteristic services are defined from the packages or plans of VoIP services, which are presented in promotional messages on operator websites. Examples are:

Skype.com: Unlimited calls, Landline calls, Mobile calls, Landline and mobile calls, credit payment, subscription payment.

aglow.sg: international calls, traditional analog phone lines, mobile phone line.

Skype.pchome: Unlimited world, Taiwan landline calls, Taiwan mobile calls, Unlimited United States calls, China calls.

Most of the VoIP operators offer similar service plans with various features. This research uses Skype as a representative VoIP service, on account of its current popularity in voice networking. Skype enables international calling through any landline or mobile phone, including received calls and voicemail messages. The main function of Skype is to support long distance and local calling by consumers.

2.2 Logistic model

Many techniques have been proposed for analyzing the characteristics or attributes of telecommunication products [2,3,4]. However, comparatively few studies in the VoIP service literature have related service characteristics to VoIP consumer choice. Here, we evaluate a VoIP transaction dataset by Bayesian logistic regression, and compare the results with those estimated by ML. ML is a preferred analysis technique because its estimators are consistent, sufficient, and efficient [5]. The parameters estimated by ML and Bayesian regression are compared in terms of their predictive accuracy in telecommunication sectors.

3. Research Methodology

In our methodology, the variables are VoIP service characteristics. We model the relationship between these service characteristics and Global Calling service choice.

3.1 Data

This data are e-retailer sales transaction data on prepaid cards sold in Taiwan. The dataset includes information on retail choice by consumers who had purchased Skype prepaid cards and had recommended used services in blogs. Many Taiwanese customers purchased from two e-retailers on the Yahoo (yahoo.com.tw) and Ruten (ruten.com) sites. The dataset comprised 573 individual records logged in 28 months between 2011 and 2013 of the service plan. We identified 9 service characteristics among 8 VoIP service plans from messages promoting prepaid cards on e-retailers' websites. These are reported in Table 1.

In summary, the data consists of 9 service characteristics as independent variables for 8 VoIP packages. The data are categorized as Global Calling or non-Global Calling as dependent variables. These are defined as the voice characteristics which influence consumer choice. Data collection is implemented from the service features advertised in the prepaid card messages. We selected 8 service plans which could be purchased at the time of data collection.

3.2 Logistic regression approach

Logistic regression is used to predict a discrete outcome. The dependent variable is dichotomous, such as presence/absence or success/failure. When the independent variables are categorical or mixed continuous/categorical, or when the dependent variable is dichotomous, logistic regression is the preferred analysis technique.

Table 1 Variables and descriptions of VoIP service characteristics used in the analysis

Variables	Definition	Mean	S.D.
<i>GLOBAL</i>	Plan call listing included call to landline and mobile for 10 countries (Canada, China, Guam, Hong Kong SAR, Puerto Rico, Singapore, Thailand, United States) and call to landline for 65 countries (available at www.Skyp.com/en/rates/)	0.430	0.495
<i>TW_JP_KR</i>	Plan call listing included call to Taiwan, Japan and Korea	0.070	0.252
<i>US_CA</i>	Plan call listing included call to USA and Canada	0.010	0.072
<i>TW_L_UN</i>	Plan call through Taiwan landline unlimited included	0.790	0.407
<i>TW_L</i>	Plan call through Taiwan landline included	0.990	0.110
<i>TW_M</i>	Plan call through Taiwan mobile included	0.482	0.501
<i>ONE_M</i>	price prepaid per one month (\$TWD)	273.770	95.303
<i>THREE_M</i>	price prepaid per three months (\$TWD)	733.170	279.707
<i>ONE_YEAR</i>	price prepaid per one year (\$TWD)	2625.630	837.394
<i>OLD_CUS</i>	Plan for old consumers	0.090	0.290

1.000 Taiwan Dollar (\$TWD) = 0.033 \$ US Dollar

In logistic regression, the relationship between the predictor variable and dependent variables is non-linear. The logistic regression function is the logarithmic transformation of the probability, and is defined as

$$\theta = \frac{e^{(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}}{1 + e^{(\alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}} \quad (1)$$

where x_i are the predictors, α is a constant and, β is the coefficient of the predictor variables.

This transformation yields a number, called logit $P(X)$, for each test case with input variable X . The logistic regression model linearly fits the log odds to the variables [6].

An alternative from of the logistic regression equation is given as

$$\begin{aligned} \log \left[\frac{\theta}{1 - \theta} \right] &= \text{logit } P(X) \\ &= \alpha + \sum_{i=1}^k \beta_i X_i \\ &= \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \end{aligned} \quad (2)$$

This study compares the two approaches of ML estimation and Bayesian estimation.

The first approach is a classical statistical method of parameter estimation. ML estimates the unknown parameters in a logistic regression model [7]. ML maximizes the log likelihood, defined as the likelihood (odds) that the observed value of the dependent variable will be predicted from the observed values of the independents. The results of logistic regression are expressed in terms of odds rather than probabilities because the odds provide a pure summary statistic of the partial effect of a given predictor, controlling for the predictors in the logistic regression.

The second approach is a Bayesian statistical method of parameter estimation. The Bayesian regression model uses the Markov Chain Monte Carlo (MCMC) approach which is simple and accurate method [8]. MCMC involves three steps: (a) formulating prior probability distributions of targeted parameters; (b) specifying the likelihood function; and (c) MCMC sampling of the posterior probability distributions. In the Bayesian approach, parameters are considered random and a joint probability model for both data and parameters is required. This difficulty is most easily circumvented by proposing an informative prior of small precision, to avoid potential bias introduced by subjective beliefs [9]. Because the joint posterior distribution of the parameters of the proposed models is analytically intractable, the point and interval estimates of the parameters are obtained by simulation-based MCMC methods [10].

After estimating a suitable prior distribution of the model parameters, the model runs simultaneously in WinBUGS. The parameters are defined from the data by a set of vague normal priors. A burn-in of 5,000 iterations is permitted, followed by 10,000 iterations in which the calculated intercept and coefficients are stored. Convergence was achieved after 5,000 iterations and the posterior distributions of model parameters were summarized using descriptive statistics. ML estimator of the model parameters was obtained in R, using the Rcmdr package introduced and developed by John Fox [11]. Comparison of the obtained results was also implemented in this package.

In this study, the dependent variable is Global Calling—*Global*—, which takes value 1 if the consumer has used the Global Calling service and 0 otherwise. The independent variables are *TW_JP_KR*, *US_CA*, *TW_L_UN*, *TW_L*, *TW_M*, and *OLD_CUS* is a dummy indicator variable. The three cost-per-period levels, *ONE_M*, *THREE_M*, and *ONE_YEAR*, are continuous variables. The analysis attempts to highlight the relationship between the VoIP service characteristics and Global Calling choice. A logistic regression model can be written in terms of the odds of an event occurring. These odds are defined as the ratio of the probability that an event will occur and the probability that it will not. The coefficients indicate the effect of individual explanatory variables on the logarithm of the odds. In terms of the variables used in this study, the logistic regression equation is

$$Y_i = \alpha + \beta_1(TW_JP_KR) + \beta_2(US_CA) + \beta_3(TW_L_UN) + \beta_4(TW_L) + \beta_5(TW_M) + \beta_6(ONE_M) + \beta_7(THREE_M) + \beta_7(ONE_YEAR) + \beta_8(OLD_CUS) + \varepsilon_i \quad (3)$$

where Y_i is the log odds of Global Calling for the i th observation and ε_i is its error term. As shown in Eq. (3), the probability of Global Calling is affected by the difference between Global Calling and its absence. In this study, the parameters are estimated by both Bayesian and ML methods.

4. Results

The results are summarized in Tables 2 and 3. The fit models of the two approaches are assessed and the estimates are reported. Our main interest lies in analyzing the relationship between the presence/absence of Global Calling in VoIP service plans and the various characteristics of the service.

4.1 Model fitting

The model is based on Bayesian and ML estimation. The eight predictors are the characteristics of the VoIP service extracted from the promotional messages of prepaid cards posted on company websites. We compared the model with respect of Akaike's Information Criterion [12] and Deviance Information Criterion [13]. The DIC value reported by Bayesian analysis was 204.8, while the AIC reported by ML was 213.89. The smaller DIC value indicates that the Bayesian method provides a better fit to the data.

4.2 Variable effects

The estimates output by the two approaches (Bayesian logistic regression and ML) are summarized in Tables 2 and 3. The results involve eight predictors and the outcome is the probability of *GLOBAL*.

In the Bayesian method, the posterior estimations of beta coefficients are obtained as the means of simulated betas from different numbers of samples after confirming convergence. On the other hand, the ML estimates of beta coefficients are obtained by iterating the Newton–Raphson method on the data. In both cases, the results are most conveniently explained in terms of the odds ratio (the exponential of the betas coefficients). The respective posterior credible intervals (or confidence intervals) can be readily obtained in the usual manner.

The variables considered are dummies, discrete or continuous variables. The odd ratio of a discrete variable can be interpreted as the change in relative risk of a certain outcome associated with a change in the independent variable. Values below 1 imply a negative association and their magnitude indicates their contribution to the probability of the outcome. The odds ratios for the continuous variables, on the other hand, can be interpreted as elasticities; that is, the odds ratio indicates the extent to which a one per cent change in an independent variable affects the relative risk of the outcome. As for the discrete variables, values below 1 imply negative elasticities. This study presents the odd ratio analysis of the discrete variables in *GLOBAL* and non *GLOBAL*.

The Bayesian analysis indicates that five variables are significantly correlated with Global Calling. Two variables, *ONE_M* and *TW_L*, make significantly positive contributions. On the other hand, ML analysis reveals only two variables that contribute significantly to Global Calling choice, of which one (*ONE_M*) makes a positive contribution. *THREE_M* exerts a significantly negative effect on the service.

The odds ratio, together with the respective 95% posterior credible intervals and 95% confidence intervals, are also presented in Tables 2 and 3. Similar results are obtained for both Bayesian and ML estimation methods. The *ONE_M* variable is the characteristic yielding the highest number of sales. *THREE_M* and *ONE_YEAR* are variables of higher cost than *ONE_M* and generate fewer sales than *ONE_M*. From these results, it is evident that consumers who select higher-cost services are less likely to intensify the service. The results are supported by both Bayesian and ML at the 5% significance level.

The *TW_L* variable exerts a positive impact on Global Calling. In the real market, if an operator enables the *TW_L* characteristic, the volume of sales is likely to

applicability of the Bayesian approach to sample observations.

Table 2: Bayesian Estimates in Logistic model

Parameter	mean	2.50%	97.50%	Odd Ratio
alpha	-3.43E + 01	-44.850	-21.120	1.33E - 15
<i>ONE_M</i>	1.10E - 01	0.073	0.146	1.12E + 00 *
<i>THREE_M</i>	-2.17E - 02	-0.032	-0.014	9.79E - 01 *
<i>ONE_YEAR</i>	1.02E - 03	-0.002	0.004	1.00E + 00
<i>TW_L</i>	2.42E + 01	8.336	35.900	3.20E + 11 *
<i>TW_M</i>	-1.00E + 02	-132.100	-70.770	2.76E - 44 *
<i>TW_L_UN</i>	-6.21E + 00	-15.200	-0.013	2.01E - 03 *
<i>US_CA</i>	-1.31E + 01	-64.830	22.900	2.02E - 06
<i>TW_JP_KR</i>	-4.93E + 00	-63.740	44.250	7.21E - 03
<i>OLD_CUS</i>	8.06E - 01	-1.216	2.892	2.24E + 00

Signif. Code: '**'

Table 3: Maximum Likelihood Estimates in Logistic model

variable	Estimate	Std. Error	Pr(> z)	Odd Ratio
Intercept	-4.41E + 01	7.44E + 03	0.995	7.33E - 20
<i>ONE_M</i>	1.49E - 01	2.68E - 02	3.10E - 08	1.16E + 00 *
<i>THREE_M</i>	-2.41E - 02	5.46E - 03	9.92E - 06	9.76E - 01 *
<i>ONE_YEAR</i>	-1.54E - 03	1.81E - 03	0.393	9.98E - 01
<i>TW_L</i>	4.52E + 01	1.24E + 04	0.997	4.22E + 19
<i>TW_M</i>	-7.54E + 01	9.93E + 03	0.994	1.78E - 33
<i>TW_L_UN</i>	-1.94E + 01	9.88E + 03	0.998	3.61E - 09
<i>US_CA</i>	1.16E + 01	1.27E + 04	0.999	1.09E + 05
<i>TW_JP_KR</i>	-2.71E + 01	2.91E + 04	0.999	1.65E - 12
<i>OLD_CUS</i>	1.03E + 00	1.59E + 00	0.517	2.79E + 00

Signif. Code: '**'

increase. Therefore, the operator is likely to design service plans that support international calling. These conclusions can be drawn only from the Bayesian analysis; ML estimation did not identify a similar impact of the *TW_L* variable.

A primary goal of the present study is to compare the cause-effect estimates from ML and Bayesian analysis. Interestingly, when the two methods were applied to Global Calling data, they yielded different results. The variables *ONE_M*, *THREE_M*, *TW_L*, *TW_M* and *TW_L_UN* showed a cause-effect relationship under Bayesian estimation, while ML estimation implicated just two variables, *ONE_M* and *THREE_M*, in the cause-effect of VoIP services. The results also demonstrate the

5. Conclusions

The purpose of this study was to predict the important variables that influence Global Calling choice probability in the Taiwanese local call market. To this end, we identified the VoIP characteristics that most strongly affect Global and non-Global Calling by logistic regression analysis using Bayesian and ML methods. The methods were compared in terms of their AIC and DIC statistics. The results show that *ONE_M* and *TW_L* make significantly positive contributions to *GLOBAL* choice, while *TW_M*, *TW_L_UN* and *ONE_YEAR* exert a negative influence. This study has also identified the Taiwan landline service package as the greatest positively contributing factor.

The Bayesian method provided a better fit to the data than the ML method, and identified a larger number of significant factors. Similar statistical studies will assist operators in elucidating VoIP service patterns. Furthermore, researchers can identify interesting transaction data that will assist theoretical development in communication technology.

The empirical results of this study have implications for Internet network operators. The main factors predicted to influence customer choice in Global Calling services depend on the adopted analysis approach. We conclude that the Taiwanese local call market should retain the landline feature in its service package.

Acknowledgments

Special acknowledgment is extended to Charnarin for assistance with data collection and for the many valuable discussions. We also acknowledge Ryan Chen for motivating this study, and who instigated useful discussions on the need to identify the characteristics of VoIP services from commercial Taiwanese websites.

References

- [1] ITU, "Internet continues to grow Worldwide to reach 2.3 billion mark late June 2012", available at: <http://www.itu.int>, accessed 5 May 2013.
- [2] F.m Tseng, "Forecasting the Taiwan customer market for Internet telephony", Journal of the Chinese Institute of Industrial Engineers, Vol. 22, No. 2, 2005, pp. 93-105.
- [3] M. L. Zubey, W. Wagner, and J. R. Otto, "A conjoint analysis of voice over IP attributes", Internet Research, Vol. 12, No. 1, 2002, pp. 7-15.
- [4] G. Cecere, and N. Corrocher, "The usage of VoIP services and other communication services: An empirical analysis of

- Italian consumers”, *Technological Forecasting and Social Change*, Vol. 79, No. 3, 2012, pp. 570-578.
- [5] M.E.B. Akiva, and S. R. Lerman, “Discrete choice analysis: theory and application to predict travel demand”, MIT press, Vol. 9, 1998.
- [6] B. W. Lindgren, *Statistic theory*, Macmillan College, 1976.
- [7] S.E. Fienberg, *The analysis of cross-classified categorical data*: Springer, 2007.
- [8] R. S. Akhter, S. Hussain and U. R. S. Habib, “MCMC simulation of GARCH model to forecast network traffic boad”, *International Journal of Computer Science Issues*, Vol. 9, No.2, 2012, pp. 277-284.
- [9] R. Haylock, and A. O’Hagan, “On inference for outputs of computationally expensive algorithms with uncertainty on the inputs”, *Bayesian statistics*, Vol. 5, 1996, pp. 629-637.
- [10] L. Tierney, “Markov chains for exploring posterior distributions”, *The Annuals of Statistics*, 1994, pp. 1701-1728.
- [11] J. Fox, “Getting Started With the R Commander: A Basic-Statistics Graphical User Interface to R”, *Journal of Statistical Software*, Vol. 14, No. 9, 2005, pp. 1-42.
- [12] H. Akaike, “Information theory and an extension of the maximum likelihood principle”, *Selected Paper of Hirotugu Akaike*, 1998, pp. 199-213.
- [13] D.J. Spiegelhalter, N. G. Best, B. P. Carlin et al., “Bayesian measures of model complexity and fit”, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 64, No. 4, 2002, pp. 583-639.

Thanyalak Mueangkhot received her M.B.A in Faculty of Liberal Arts and Management of Kasetsart University, SakonNakhon campus, Thailand, in 2005. She is now a graduate in Department of Business Administration of National Central University. Her main research interest is service marketing, green marketing, and Internet Technologies.

Julian Ming-Sung Cheng is an Associate Professor of Marketing in the Department of Business Administration, National Central University, Taiwan. His current research interests include marketing channels, international branding, global marketing and internet marketing. His papers have appeared in *Journal of International Marketing*, *Industrial Marketing Management*, *Journal of Business and Industrial Marketing*, *Journal of Advertising Research*, *International Journal of Market Research*, *Journal of Product & Brand Management*, *Journal of Marketing Channels*, *Asia Pacific Journal of Marketing and Logistics* and so on.

Chaknarin Kongcharoen received his M.Eng in Computer Engineering in Faculty of Engineering, Suranaree University of Technology, Thailand, in 2006. He is now a graduate in Department of Computer Science & Information Engineering, National Central University. His main research interest in data mining, web mining, software metric and e-learning. His papers have appeared in *KMITL Science and Technology Journal*.

EVAM-MAC: An Event Based Medium Access Control for Wireless sensor Networks with Multihop Support

Zaher Merhi¹, Mohamed Elgamel², Samih Abdul-Nabi¹,
Amin Haj-Ali¹, and Magdy Bayoumi³

¹Department of Computer and Communication Engineering
Lebanese International University
Beirut, Lebanon
{zaher.merhi, samih.abdunabi, amin.hajali}@liu.edu.lb

²Arab Academy for Science
Technology and Maritime Transport Alexandria, Egypt
cshaheen@hotmail.com

³Center for Advanced Computer Studies
University of Louisiana at Lafayette
LA, USA
mab@cacs.louisiana.edu

Abstract

As wireless sensor network applications are becoming more complex, the need for a versatile medium access control that is able to deliver high data rates is essential. In event-based systems, sensor nodes spend most of their time at sleep state waiting for an event to occur. When an event is detected, sensor nodes experience a short abrupt period of high data contention where the data packets are large in size. As the number of sensor nodes per-hop increases, the contention generated will lower the throughput, increase latency and deteriorate the application's performance. EVAM-MAC is a medium access control that is tailored specifically for event-based systems. EVAM-MAC shifts the contention generated by multiple sensor nodes trying to deliver the collected measurements to the control phase. Furthermore, it arranges data transfer in a contention free environment by dynamically creating a TDMA-like schedule without global synchronization or global slot assignments. EVAM-MAC offers a platform of configurable operations that can be programmed prior to deployment. The protocol is simulated with NS2 and compared with an implementation of S-MAC and 802.11 for multiple scenarios where EVAM-MAC presented its superiority in throughput, Latency and Energy con-

sumption.

Keywords : *Clustering Algorithms, Event Detection, Medium Access Control, Wireless Sensor Networks*

1 Introduction

Medium access control protocols (MAC) are considered an integral part of any wireless sensor network (WSN) application. These applications are usually constrained by both power and time. Typically, WSN are battery operated where replacing and/or charging the battery is not always feasible. Thus, it is crucial for the application longevity to preserve power consumption. Furthermore, real time applications such as target tracking, cannot tolerate latency in data delivery. Dropped packets due to high data contention wastes both power and time. Initially, the trend with MAC protocols for wireless sensor networks was to trade for latency with energy consumption by duty cycling the radio transceiver [2] [3]. This suited monitoring application since wireless sensor devices transfer small amount of information every periodic timeframe. On the other hand, event based applications spend most of their time at sleep state waiting

for an event to occur. Once an event is detected, a short abrupt period of high data contention is exhibited due to multiple nodes contending for the channel. Thus, trading energy for latency is no longer a valid choice [1].

High throughput applications usually employ TDMA schemes. However, they are not efficient for wireless sensor networks [5], [6]. In TDMA schemes, data transmission is regulated by assigning a unique slot for each sensor node in the two-hop neighborhood. In event-based systems, only sensors detecting the event should participate in data transmission. Thus, time slots are wasted where sensors not wishing to transmit are forced to sleep. Slot reuse is employed by hybrid TDMA/CSMA schemes as in [7]. However, the overhead in creating and maintaining a schedule across multiple sensors at multiple hops consumes significant amount of energy. Although this overhead only occurs at deployment the cost of updating the slot distribution across the network for nodes joining and leaving the network (which is frequent in wireless sensor networks) is significant and cannot be neglected. Moreover, these schemes also employ global time synchronization that incur even more overhead in packet exchange and in maintaining clock drift corrections over time.

On the other hand, wireless sensor network is an application centric device. MAC protocols should offer a set of policies and procedures that can be configured to optimize event-based systems according to the applications' designers need. Moreover, event-based systems exhibits certain characteristics that the MAC protocol should take into consideration to achieve an optimized design. For an example, dynamic clustering is very desirable in event-based systems as only sensors detecting that event should wake up and deliver measurements to a local cluster head. Furthermore, it is more efficient to aggregate data at local cluster heads rather than having each node transmitting its measurements throughout the network. Static clustering is not efficient for wireless sensor applications since nodes are consistently joining and leaving the network. However, as sensor node count increases per cluster, the contention is heightened as multiple sensors try to deliver their recorded measurements to the cluster head. Since measurements are highly corrupted by noise in wireless sensor networks, prioritizing data transfer according to the received signal strength (RSS) of the event restrains the system from overshooting and insures quality of service [9]. On the other hand, measurements recorded from the captured event are highly correlated which make it sometimes useless to send more than a certain number of measurements according to a certain policy. For an

example, in a 10-node network scenario, an application could be satisfied with only 5 high quality readings (i.e. high RSS). Transmitting all 10 readings will increase contention and waste both bandwidth and energy. A mechanism should be devised in order to instruct a subset of sensors belonging to a cluster to transmit according to the quality of their readings.

Furthermore, optimizations can be achieved by tailoring the MAC protocol to a pattern of communications that the application requires. For example, distributed signal processing application usually share the readings recorded by the sensor nodes that are in the one-hop neighborhood for processing via broadcast. Afterwards the result is unicasted to the cluster head for aggregation. Some applications require several rounds of such. The control packet overhead (RTS-CTS-ACK) can be minimized if the sensor nodes participating in the event register their interest. Thus, the MAC protocol can be configured to follow the patterns required by the application by dynamically constructing an efficient schedule. Lowering the control overhead will increase throughput and lower energy consumption by restraining the sensor nodes from contending for the channel each time a communication sequence is started for the same event when detected.

EVAM-MAC is an event based medium access control that addresses all the problems listed above. When an event is detected EVAM-MAC will dynamically create a virtual cluster containing only nodes that detected the event. The cluster is constructed by introducing a control gap that registers sensor nodes that have recorded viable data. A Schedule will be created for this cluster that is unique across the two-hop neighborhood. Thus, each node will have a contention free slot to send their data. EVAM-MAC is the extension of our work presented in [8]. This article is organized as follows: Section II discusses the related work, Section III presents EVAM-MAC protocol design, Section IV discusses the simulation experiments and results, and Section V is the conclusion and future work.

2 Related Work

Most medium access control for wireless sensor networks lowers energy consumption by decreasing the duty cycle of the radio. Having a low duty indicates that the radio is spending most of the time at sleep state. However, this has the effect of lowering the throughput. In Event-based systems the sensor nodes spend most of their time at sleep state waiting for an event to occur. Once the event is detected, a short

abrupt period of high data contention follows.

S-MAC [2] and T-MAC [3] are scheduled based protocols that preserve energy by duty cycling the radio. These MAC protocols trade energy with latency. However, in event-based systems it is desirable to have a real time response. S-MAC synchronizes nodes using a SYNC messages that assigns a schedule to each node. The overhead produced by the SYNC messages introduces huge latencies as the one hop neighborhood grows in size. T-MAC [3] improves on S-MAC by reducing the idle time spent using a time-out scheme. All communications are moved to the beginning of the frame which set the nodes into a fierce mode for acquiring the channel. However, it still suffers from the shortcomings of S-MAC.

Hybrid protocols such as the Z-MAC [7] and F-MAC [12], solves the problem of high data contention by incorporating both CSMA (Carrier Sense Multiple Access) and TDMA (Time Division Multiple Access) techniques. In Z-MAC, CSMA is used in low contention and TDMA is used in high contention. it uses DRAND [13] which assigns unique slots for the TDMA scheme in the two-hop neighborhood which is performed at deployment. This operation is a resource intensive task that includes neighborhood discovery, the operations of DRAND, and global synchronization. However, the overhead exhibited in nodes joining and leaving the network is large and cannot be neglected. Global synchronization also induces an overhead in maintaining coherent timing information that deteriorates the system performance.

Leach [14], [15] protocol groups sensor nodes into cluster where each node reports to the local cluster head. Data aggregation takes place at the local cluster head where aggregated data is sent directly to the base station. Leach operations are divided into rounds where each round consists of a setup and steady state phase. The disadvantage of Leach is that all nodes are required to be in communication reach of a base station which is impractical for wireless sensor network applications. Moreover, the number of sensors nodes and clusters formed are required to be known a priori which limits the flexibility that characterizes wireless sensor networks.

In BMA [4], the cluster head forms a schedule by arranging data transfer at each round. The cluster head accepts requests for data transfer and assigns a slot for each node wishing to transmit. Each round of data transfer is divided into contention, data transmission and idle period. Cluster formation is done by using Leach protocol [14]. The drawback of this protocol is that it permits centralized control where the cluster head is responsible for maintaining a schedule across nodes belonging to the same cluster. However,

in wireless sensor networks it is more efficient to have a decentralized control which eliminates single point of failure. Moreover, nodes that have no data to send waste time slots in the contention period where idle listening and overhearing occurs. Furthermore, each session has a fixed time for its completeness. Thus, if few nodes wish to send data, idle time is wasted there which increase latency. Moreover, in dense deployment, if more nodes wish to send data, sessions may overlap and channel contention may occur in the data transmission period especially if they join before the cluster head set up its candidate nodes.

B-MAC [16] and X-MAC [17] utilizes CCA (clear channel assessment) to assess channel clarity and LPL (low power listening) for duty cycling the radio. Channel activity is checked by measuring the signal strength samples when the channel is assumed to be free. Nodes communicate with each other by sending a long preamble that is twice the check interval of the receiving nodes. Thus, nodes at sleep state have enough to time to wake up and detect the transmission. SCP-MAC [18] focuses on lowering the duty cycle by combining scheduling and channel polling. However, similar to previous protocols, they suffer from the problem of poor operation under abrupt high contention traffic especially in multihop scenarios.

AS-MAC [21] utilizes hello packets to build a neighborhood table. The receiver of the data packet follows a periodic schedule that the sender is aware of. Based on this schedule Nodes wishing to send data packets will wake up accordingly. AS-MAC uses static clustering with is inefficient in event-based systems. Likewise, RI-MAC [19] also relies on the receiver to announce its readiness to receive Data transmission by sending a beacon frame. All nodes follow a periodic schedule to check for pending data transmission. When a beacon is received, nodes with pending data start transmission. PW-MAC [20] enhances RI-MAC by introducing a predictive wake up schedule on the sender side. The sender node will predict the approximate time of the receivers wake-up schedule and transmit data based on that time. Although these MAC protocols present an enhancement on synchronous scheduled protocols, they still are not optimized for event-based systems. EVAM-MAC is specifically designed to address the needs of event-based systems where sensor node count per-hop is high. EVAM-MAC dynamically creates a contention free schedule each time an event is detected. EVAM-MAC achieves high throughput, maintains acceptable energy consumption levels, and control packet overhead due to the reduced number of dropped data packets.

3 EVAM-MAC Protocol Design

The main principle behind EVAM-MAC is to shift contention generated from multiple sensor nodes (in the two-hop neighborhood) transmitting the collected measurement from the data phase to the control phase. In event-based applications for WSN such as target tracking, sensor nodes transmit simultaneously large amounts of data. Thus, dropped packets due to the high contention generated, have a significant effect on performance. Having an optimized contention free schedule is crucial for real time operation and energy savings. EVAM-MAC creates a TDMA-like schedule dynamically, and in a decentralized fashion. Furthermore, it eliminates the use of global synchronization that has a high overhead to maintain and substitutes for it with local synchronization. EVAM-MAC does not assume any infrastructure constraints and does not require a setup phase as in hybrid protocols which facilitates nodes joining and leaving the network. Moreover, it prioritized the schedule based on received signal strength (RSS) reading of the collected data by giving high priority for sensor nodes with high RSS. This way sensor nodes with reliable data can be given a higher privilege than those with lower ones which otherwise can deteriorate the convergence of the application.

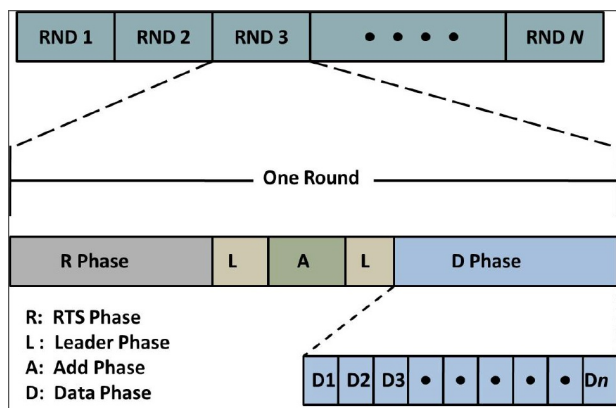


Figure 1: Round operations

EVAM-MAC's operation is divided into rounds as shown in Fig. 1. When an event is detected, only sensor nodes detecting that event initiate a round. Each round is dedicated for a single event where subsequent events are buffered for future rounds. The purpose of a round is to fulfill the application layer requirements for the successful delivery of the infor-

mation for processing. This includes the communication pattern dictated by the application, such as, a broadcast of the collected data and then a unicast of the processed data to a dedicated node. EVAM-MAC adopts a modified version of B-MACs CCA and LPL techniques to assess channel activity and to duty cycle the radio. Instead of using a long preamble for LPL phase, EVAM-MAC replaces it with a train of packets to accommodate the CC2420 radio that is available on most wireless sensor nodes. Once an activity is detected during LPL phase the sensor nodes stays awake for the entire round. As shown in Fig. 1, each round is divided into sub-rounds where each one has a specific role.

3.1 Neighborhood Discovery

The main purpose of the RTS Sub-round operation is to allow nodes detecting the same event to be included in the contention-free schedule. Once an event is detected by a group of sensor nodes that are in the communication reach of each other the RTS sub-round timer is initiated as shown in Fig. 2. EVAM-MAC assumes that sensor nodes detecting the same event are in the 1-hop neighborhood of each other. This assumption is considered reasonable since the communication range is usually much larger than the sensing range for most applications. Only nodes detecting the event will start broadcasting RTS packets to notify all nodes that a round is started. Each node checks for channel activity (using CCA) before transmitting its RTS packet. If the channel is busy, it backs off and tries again. Nodes periodically send their RTS packets to notify other nodes about the event. No acknowledgement is used for the RTS packets since each node detecting the event will continuously broadcast the RTS packet regardless of any successful reception as shown in Fig. 2. The reason behind this implementation is that we don't know a priori the number of nodes in the communication reach of the transmitting node or the number of nodes detecting the event. The RTS sub-round timer value is an application specific parameter that is provided according to the approximate sensor count and application needs. Higher values of the RTS sub-round timer will give more chances for nodes to be included in the schedule. The reason behind this is that as the sensor count increases, collisions also increases and some nodes will not be able to successfully broadcast their packets before the round timer expires. However, care must be taken when choosing the RTS sub-round timer value since large values will incur unnecessary overhead and with low values nodes may not be able to participate in the schedule.

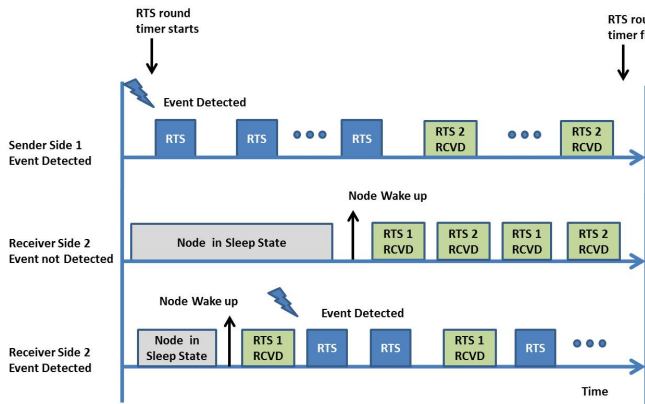


Figure 2: RTS Round operations

The RTS round phase is responsible for choosing the leader node, synchronizing neighboring nodes, dynamically creating the cluster and constructing a preliminary schedule. All of these operations are performed simultaneously in the RTS sub-round. The RTS packet is piggybacked with the following attributes:

- RSS: The received signal strength of the detected event (8-bits).
- Leader_Node: The node with the lowest time to detect the event (8-bits).
- Recv_Addr: The receiver address of the data packet (8-bits).
- Rts_Leader: Indicates if the sender node is a leader (1-bit), used for multi-hop resolution.
- TTE (Time to expire): The time required to finish the RTS sub-round timer (16-bits).
- ID: Node ID (8-bits)

3.1.1 Leader Selection and Time Synchronization

Since each node detecting the event starts its RTS sub-round timer independently, time synchronization is required for the next sub-rounds. When the schedule is created, each node detecting the event must have a slot that is synchronized and unique across the 2-hop neighborhood in the DATA sub-round. Otherwise, overlapping of data transmissions may occur which defies the whole purpose of EVAM-MAC (i.e. having a contention free channel for DATA transmissions).

EVAM-MAC implements only local synchronization each time an event is detected or a round is initiated. For the RTS sub-round, each node is faced with

two options, if it has detected the event or not. A node detecting the event will have its RTS sub-round timer initiated and will be sending out RTS packets. Moreover, it will assume leadership of this round. Leadership status has no advantages in terms of processing or capabilities. It is just a privilege granted to a node so that it can maintain synchronization and scheduling. Nodes detecting the event dynamically elect the Leader node. The criteria used to choose the Leader node is the node with the earliest time to detect the event. This value is computed by determining the remaining time to expire (TTE) of the RTS sub-round timer which is piggybacked with the RTS packet. Sensor nodes not detecting the event in the two-hop neighborhood will be only receiving the RTS packets. Moreover, these nodes will update their RTS sub-round timer to be synchronized with other nodes. This step is crucial for nodes detecting a subsequent event not to interfere with an ongoing round. Furthermore, these nodes are forced to sleep for the entire DATA sub-round if they are not the intended receiver of the data to be sent. When a node receives an RTS packet it checks first if its own RTS sub-round timer has already started. If not, it will update the leader node to the senders ID and start its RTS sub-round timer. The value of the timer is set in accordance to the remaining time of the RTS sub-round timer of the sender. If the RTS sub-round timer has already started, it will check if it should remain the leader. If so, it will add the senders ID to the Schedule and discard timing information. If the node is not the earliest to detect the event, it will update the leader node and its RTS sub-round timer. The details of the RTS sub-round is presented in Algorithm 1

To illustrate the synchronization phase of the RTS sub-round operation, consider a 3 node topology that is shown in Fig. 2. The details of the operations is presented in Fig. 3. In this topology nodes 1 and 3 have detected the event and node 2 is only monitoring the round. Node 1 is the first node to detect the event and consequently it will be the leader of this round. As shown in Fig. 3, the RTS sub-round timer of node 2 is started with the remaining portion of the Leaders RTS sub-round timer minus the time to receive an RTS packet (TTR). TTR also include any processing delays exhibited at the receiver and the sender. These delays can be approximated based on the sensor nodes platform. Node 3 will have to relinquish its leadership status and it will update its timer based on the value received from node 1 minus TTR. Nodes 1 and 3 will keep on broadcasting RTS packets until the RTS sub-round timer expires where both carry the same leader and TTE informa-

tion. All nodes receiving RTS packets will have their timer value expiring at the same time and these nodes will be ready for the subsequent round.

Algorithm 1 EVAM-MAC RTS sub-round operations

```

Duty_Cycle(Node(i))
if EventDetected then
    if RTS_STARTED == FALSE then
        RTS_Timer.start(RTS_ROUND)
        Leader_Node = i
        createRTSPacket(RSS, LEADER_Node,
            Rts_Leader, Recv_Addr, TTE, ID)
        if RTS_Timer.Expired() == FALSE then
            sendRTSTrain()
        end if
    else
        createRTSPacket(RSS, LEADER_Node,
            Rts_Leader, Recv_Addr, TTE, ID)
        if RTS_Timer.Expired() == FALSE then
            sendRTSTrain()
        end if
    end if
end if
if Received(RTS_PACKET) then
    if RTS_STARTED == FALSE then
        {didnot detect the event}
        RTS_Timer.start(RTS_ROUND
            -RTS_Packet.TTE - TTR)
        {TTR: time to receive RTS packet}
        Leader_Node = RTS_Packet.Rts_Leader
        AddNodeToSchedule(RTS_Packet.ID,
            RTS_Packet.Recv_Addr, RTS_Packet.RSS)
        {arrange schedule in decreasing order of RSS, save
            destination address of the data packet}
    else
        TTE_Recv = RTS_ROUND -
            RTS_Packet.TTE - TTR
        if TTE_Recv < TTE then
            RTS_Timer.update(TTE_Recv)
            Leader_Node = RTS_Packet.Rts_Leader
            AddNodeToSchedule(RTS_Packet.ID,
                RTS_Packet.Recv_Addr, RTS_Packet.RSS)
        else
            AddNodeToSchedule(RTS_Packet.ID,
                RTS_Packet.Recv_Addr, RTS_Packet.RSS)
        end if
    end if
end if
    
```

3.1.2 Dynamic Clustering and Schedule Creation

Most event-based WSN applications are based on cluster topology. This is mainly due to that fact that redundant measurements are collected by the nodes detecting the same event. It is inefficient to broad-

cast all collected data across the network as it wastes unnecessary bandwidth and exhibits an overhead in energy consumption. Data is usually aggregated on the local cluster head before it is broadcasted to the base station. Clustering is usually static or dynamic. With static clustering, nodes arrange themselves into groups based on the communication range. On the hand, dynamic clustering is achieved by creating a group each time an event is detected. Although static clustering has a lower overhead in creation, dynamic clustering is more efficient. This is mainly because the sensing range is much smaller than the communication range in most applications. Moreover, as the event is traversing the sensor field and arriving at the boundaries of two or more clusters, nodes detecting the event will deliver their data to different cluster heads. This has the effect of increasing congestion and deteriorating the application performance. Moreover, dynamic clustering eliminates single points of failure and minimizes the overhead of nodes joining or leaving the network. In our implementation, the elected leader node acts as a cluster head every time an event is detected.

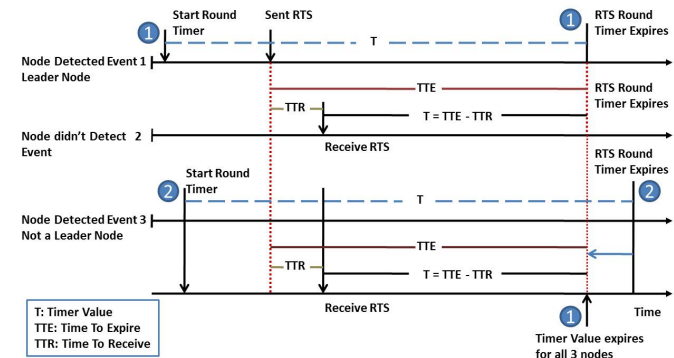


Figure 3: Locally synchronizing nodes and leader selection: All nodes are synchronized with the timer value of the leader at 1

EVAM-MAC achieves dynamic clustering in the RTS sub-round. The cluster is created by the nodes that detected the event. As shown in Algorithm 1, when a node receives an RTS packet, it adds the source address of the packet to the virtual cluster. Moreover, it saves the RSS of the event detected and the receiver address. At the end of the RTS sub-round, each node detecting the event will generate the list of nodes that are participating in the event in a decentralized fashion. The schedule for the data sub-round is created by assigning a slot for each node in decreasing order of the RSS. If two nodes have the same RSS values, lower ID nodes will win the slot. Although the leader node

is guaranteed to be unique within all nodes, the neighborhood list may have some variations from node to node depending on the level of congestion. This is mainly because not all sensor nodes will receive all RTS packets. The disambiguation in the schedule between the nodes is resolved in the Leader and Add sub-rounds.

3.2 Leader, Add and Data Sub-rounds

As shown in Fig. 1, four sub-rounds follow the RTS sub-round: Leader sub-round-1, Add sub-round, Leader sub-round-2, and the Data sub-round. The details of these four sub-rounds is presented in Algorithm 2. The purpose of the Leader sub-rounds is to resolve the discrepancies between schedules that were created independently by sensor nodes detecting the same event. This sub-round has a predefined value that is known to all sensor nodes during which all nodes are awake and only the elected leader node is allowed to transmit. The leader node will send out a Leader packet that contains its own created schedule. Each node receiving this packet will discard its created schedule and adopts the leaders schedule. The reason why all nodes create a schedule dynamically is that we dont know a priori the ID of the elected leader. The dynamic nature of EVAM-MAC offers the ability for all nodes to be a leader node provided that they have the lowest time to detect the event. This information is not resolved until the RTS sub-round timer expires.

When nodes receive the Leader packet, they will check if their ID is present in the schedule provided that they have detected the event. If their ID were not found, they have the opportunity to add it in the Add sub-round. Nodes that have their ID present in the schedule are forced to sleep for the entire value of the Add sub-round thus reducing overhearing. The Add packet contains the senders ID and the intended receiver of the data packet. Once the leader receives this packet, it will add the ID of the sender at the end of the schedule. An Acknowledgment packet is sent out to the sender to restrain the node from sending the Add packet again.

As with all sub-round timers, the add sub-round timer is used to synchronize EVAM-MAC sub-rounds. The value of the Add sub-round timer is based on statistics done on the topology of the sensor nodes. EVAM-MAC assign the value of the add sub-round timer to be 10% of the RTS sub-round timer. Although the add sub-round exhibits an overhead on EVAM-MACs round operations, its existence enhances the overall performance. Depending on the level of congestion, some sensor nodes will not have

a chance to be included in the schedule especially in dense topologies. These nodes will have to start another round for the same event which exhibits a much larger overhead than the add sub-round. Increasing the value of the RTS sub-round timer will not have the same effect since in that sub-round all sensor nodes will continuously be sending out RTS packets in a high contention environment. In the Add sub-round, only sensor nodes that are not included in the schedule send out the Add packet. Thus, nodes are given the opportunity to send out their missing packets in an environment with less contention.

At the end of the Add sub-round, all nodes will wake in preparation for the Leader sub-round-2. This phase has the same effect and operation of Leader sub-round-1. The appended schedule is sent out to all sensor nodes in the communication reach and each node updates its schedule information. After the second leader phase the data sub-round is initiated. The data sub-round is divided into slots. The slot is granted for nodes that are included in the final schedule. Each node has a unique slot that is guaranteed to be free across the two-hop neighborhood (as shown in the next section). Thus, the total value of the Data sub-round timer varies from round to round depending on the number of nodes detecting the event. Nodes that did not detect the event and are not the intended receiver of any Data packet are forced to sleep for the entire Data sub-round. During each slot only the sending node and the receiving node are awake, thus reducing idle listening. In case of a broadcast all nodes in the cluster will be awake to receive the data. As opposed to TDMA like protocols, there are no wasted slots here. Only nodes that have data to send are included in the schedule which reduces latency. The slot size is computed according to the size of the data packet. Although the medium is considered to be free, an optional ACK packet can be sent during the slot time to ensure delivery. Nodes that did not receive an ACK packet will participate in the next round to send their captured data again.

3.3 Multi-hop resolution

Multi-hop networks poses a threat on the correctness of EVAM-MAC since intersecting cluster as shown in Fig. 4 can introduce ambiguity in schedule information of each cluster. To resolve this issue, only one cluster in the two-hop neighborhood of the leader node is allowed to transmit at a time. Each time a cluster terminates its round, the remaining intersecting clusters contend for the current round at the RTS sub-round as shown later. Usually in event-based systems for wireless sensor networks, the sensing range

is much smaller than the communication range. Thus in real deployments, sensors detecting the event will yield small sparse disconnected clusters with no interference between them. Events occurring in sparse clusters have no effect on the operations of EVAM-MAC. Events occurring interchangeably (i.e. not at the same time) in intersecting clusters, as shown in Fig. 4, also have no effect on the correctness of EVAM-MAC provided one round finishes before the other. Moreover, nodes detecting an event after the RTS sub-round timer of the neighboring cluster expires, will be aware of that round and will not initiate it until the current round is finished.

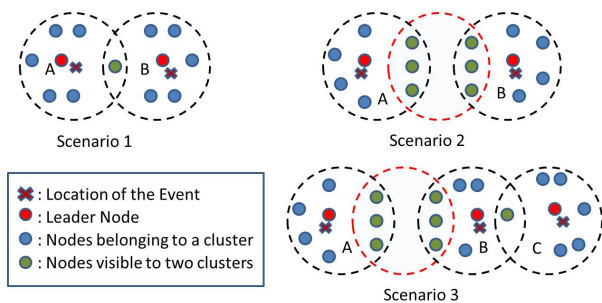


Figure 4: Two hop scenarios

As shown earlier, all nodes detecting the event will broadcast the RTS packet to all neighboring nodes. Nodes belonging to two clusters (as shown in Fig. 4, scenario one) and/or are in the communication range of some nodes in each cluster (as shown in Fig. 4, scenario two) will notify each other by the RTS packets. Consider the case shown in Fig. 4, scenario one. Assume that Cluster A detects the event before Cluster B and that the latter detects the event any time after the RTS sub-round timer of the former expires. The intersecting node will broadcast its RTS packet to all the nodes in Cluster B that is in its communication reach within the duration of the RTS sub-round timer. These nodes will start their RTS sub-round timer upon the reception of the RTS packet from the intersecting node. When the RTS sub-round timer expires, these nodes will wait for the leader node to transmit its Leader packet. However, the leader node is not in communication reach of the nodes in Cluster B. The reason behind this is that nodes that can overhear the leader node will belong to the cluster of the leader which is Cluster A. After the Leader sub-round expires, these nodes will not receive a Leader packet. They will then raise a multi hop flag and will be forced to sleep for the entire duration of the next sub-rounds. However, these nodes have no information about the duration of the Data sub-round since

the number of sensors in the schedule of cluster A is unknown to them. In order to relieve any possible interference, the amount of time assigned to the Data sub-round for nodes in Cluster B are given the maximum value which is equal to the time to transmit a DATA packet times the maximum number a cluster can attain. Although this technique introduces delays, it is more efficient than actually forcing the node to contend for the medium as it will introduce interference on the transmitting nodes in Cluster A. Moreover, it will reduce overhearing thus achieving higher energy efficiency. The maximum number any cluster can attain can be obtained before deployment by approximating the number of sensors that are in communication reach of each other. Any event occurring at cluster B after the RTS sub-round timer expires will be buffered until nodes in Cluster B are woken up. After the sleep timer expires, all nodes will repeat the whole operation again.

On the other hand, multiple events can occur simultaneously between neighboring clusters before the RTS sub-round timer expires. In this case all clusters in the communication range of each other contend to win the round. The cluster with the lowest TTE (time to expire) will win this round. If both clusters have equal TTE, the lowest ID leader node will win the round. All other clusters are forced to sleep for the entire round and defer their data transfer to the subsequent round. Consider the case shown in Fig. 4, scenario two. Following the same analogy as before, each node detecting the event will start their RTS sub-round timer and broadcast periodically the RTS packet. Let's assume that Cluster B detects the event before Cluster A. The nodes that are in communication reach of Cluster B will hear the RTS packet and update their leader and TTE information. These nodes will then broadcast their RTS packets piggybacked with the new information. However, care must be taken to choose an RTS sub-round timer value that is large enough to make this propagation possible. Upon a successful broadcast of an RTS packet of the leader node at Cluster B, any of the boundary nodes will notify Cluster A with the existence of the other round by also broadcasting a RTS packet. As discussed earlier, the cluster will shut down for the entire round for the duration of the following sub-rounds where the Data sub-round has the maximum value. Nodes belonging to Cluster A will commence their round operation after the sleep timer expires.

Algorithm 2 EVAM-MAC Leader, Add, Data sub-rounds operations

```
if RTS_Timer.Expired() == TRUE then
  Leader1_Timer.Start()
  if IsLeader() then
    createLeaderPacket(Schedule[0, 1, ..., n])
    sendLeaderPacket()
  end if
end if
if Leader1_Timer.Expired() == TRUE then
  AddTimer.start()
  if CheckNodeID(Schedule) == TRUE then
    sleep()
  else
    createAddPacket(ID, Recv_Addr)
    while ACKNotReceived() do
      sendADDPacket(LeaderNode)
    end while
  end if
end if
if Add_Timer.Expired() == TRUE then
  wakeup()
  Leader2_Timer.Start()
  if IsLeader() then
    createLeaderPacket(Schedule[0, 1, ..., n
    +appenedNodes])
    sendLeaderPacket()
  end if
end if
if Leader2_Timer.Expired() == TRUE then
  CheckSchedule()
  if CheckNodeID(Schedule) == FALSE then
    sleep()
  else
    wakeup(SlotID)
    if IsSender() then
      sendData()
    end if
  end if
end if
if Received.LeadersPkt() then
  Schedule = LeadersPkt.Schedule
else if Received.AddPkt() then
  Schedule.append(AddPkt.ID)
end if
```

In the case where we have more than two adjacent cluster as shown in Fig. 4, scenario three, both Cluster A and C are forced to sleep provided that Cluster C detected the event after Cluster B. Consider the case where Cluster A detect the event first then Cluster C and lastly Cluster B. Provided enough time is given for the RTS sub-round timer, Cluster C will also be forced to sleep. However, Cluster C can commence a round in parallel with Cluster A with no interference since Cluster B is forced to sleep. A

mechanism must be devised to notify Cluster C that it can commence its round operations. In this context, a bit (RTS_Leader) is piggybacked with the RTS packet that is set if the RTS packet was received from the actual Leader node and unset if the packet was forwarded from another node that is not the Leader node. However, if a node received a RTS packet from the Leader node once, it will always set the RTS_Leader bit disregarding any forwarded packets with the same Leader address and TTE information. Cluster B will forward the RTS packets to Cluster C with the RTS_Leader bit unset. Upon the reception of the forwarded RTS packet, Cluster C will detect that the RTS_Leader bit is unset and will disregard the Leader and TTE information and commence with its round. Thus the two Clusters (A and C) can operate in parallel. The disadvantage of this technique is that if Cluster B detects the event before Cluster C and the RTS packets from Cluster A didn't have enough time to reach Cluster C, Cluster C has no way of knowing the existence of Cluster A and assumes that its Cluster B's turn. Thus, it has to wait for the next round to operate.

3.4 EVAM-MAC Programmable features

The dynamic nature of EVAM-MAC makes it customizable for the applications needs and requirements. Programmable feature is an option that is implemented in EVAM-MAC that allows customizable Data sub-rounds. These options can be set by the application layer at deployment. The default operation of EVAM-MAC is to unicast the data packet to the elected cluster head which will forward them to the sink. However, some applications require several rounds of communications provided by the same sensor nodes detecting the event. These rounds can be a mixture of broadcast and unicast. In target localization applications, sensor nodes usually broadcast the collected measurement to the all sensor nodes detecting the target for processing. Once a location is computed, each sensor node will unicast the results back to the cluster head for aggregation. These operations can be programmed in EVAM-MAC where each round consists of 2 data sub-rounds, one for broadcast and the other for unicast. Multiple data rounds can be added depending on the applications needs. In this case the same schedule is used over again with the desired communication pattern. The overhead for multiple rounds is much lower than those of a single round with respect to the amount of Data transfer since the amount of control packet exchanged stayed the same. Thus, by lowering the control overhead

that is required to construct the data sub-rounds, and giving the application designers the freedom to choose their communication patterns and operations, optimizations can be achieved in throughput, energy consumption and control overhead.

Restraining sensor nodes transmissions is another programmable feature that is resourceful in WSN application. In dense deployments, sensor nodes collect several measurements for the same event that are highly correlated. Some applications require only a subset of sensor nodes detecting a certain type of event to transmit its collected data. For an example, consider a 15-node cluster monitoring a certain event. The application layer can be satisfied with only 5-high quality measurements. The number of measurements required and the criteria used to indicate the quality of the measurements is based on the application. EVAM-MAC arranges data schedule based on RSS of the event detected. If this feature is enabled, the Leader node will broadcast a schedule containing only the required sensor nodes disregarding the rest. Upon the reception of the Leader packet, the sensor nodes will check if their ID is present in the schedule. When they realize that they are off the schedule, they force themselves to sleep for the entire Data sub-rounds. This way we achieve higher energy conservation by prohibiting all remaining sensors from transmitting their redundant Data packets.

4 Experimentation and Simulation Results

EVAM-MAC was implemented and simulated using Network simulator (NS-2 [10]). Furthermore, EVAM-MAC is compared with an NS-2 implementation of S-MAC [2] and 802.11 [11], for one-hop, two-hop and multi-hop benchmarks. As shown in Table I, the communication bandwidth used for all benchmarks was 19.2kbs and the DATA packet payload was 60 Bytes. The RTS sub-round timer was assigned a value of 200ms for one-hop networks and 300ms for multi-hop networks since in multi-hop networks more time is required at the RTS sub-round phase to propagate the RTS packet to neighboring clusters. The Leader and Add phases are fixed to 11ms and 22ms. All sub-round values depend on the communication bandwidth. Higher bandwidth values will yield smaller sub-round values. Table II presents a comparison of the control packet sizes for all three protocols. The Leader packet of EVAM-MAC has a large value of 27 bytes since it contains the piggybacked schedule. However, the Leader packet is only broadcasted twice per round.

Three metrics were used to compare EVAM-MAC's performance with other protocols: net throughput, average energy consumption, and percentage of control packet overhead for all three benchmarks. The net throughput is computed by counting the number of DATA packets successfully received by the sink where only the payload is considered. The average energy consumption is computed by averaging the energy consumed by all nodes in the network and dividing it by the total sensor count for the duration of the simulation run. The percentage of control packet overhead is computed by averaging the number of control packet bytes sent for all nodes in the network with respect to the number of data bytes successfully received by the sink. For all three metrics, the setup phase of S-MAC was not included in the computation since it is an overhead that occurs only at deployment. However, the sync messages used after the nodes are synchronized are considered in the computations.

Table 1: Default Parameters for EVAM-MAC protocol

RTS-subround one-hop	200 ms
RTS-subround multi-hop	300 ms
Leader sub-round	11 ms
ADD sub-round	22 ms
DATA sub-round	Variable
Data Payload	60 Bytes
Communication Bandwidth	19.2 kbs

Table 2: Packet Size Information for Differnet Protocols

Type	EVAM-MAC	S-MAC	802.11
RTS	16 Bytes	10 Bytes	44 Bytes
CTS	NA	10 Bytes	38 Bytes
Leader	27 Bytes	NA	NA
ADD	5 Bytes	NA	NA
Sync	NA	10 Bytes	NA
ACK	5 Bytes	10 Bytes	38 Bytes

4.1 One-hop Benchmarks

The one-hop benchmarks consist of 20 nodes which are all in the communication range of each other and of a sink. Two applications were employed for this evaluation. The first application is a UDP (User Datagram Protocol) agent with a CBR (constant bit

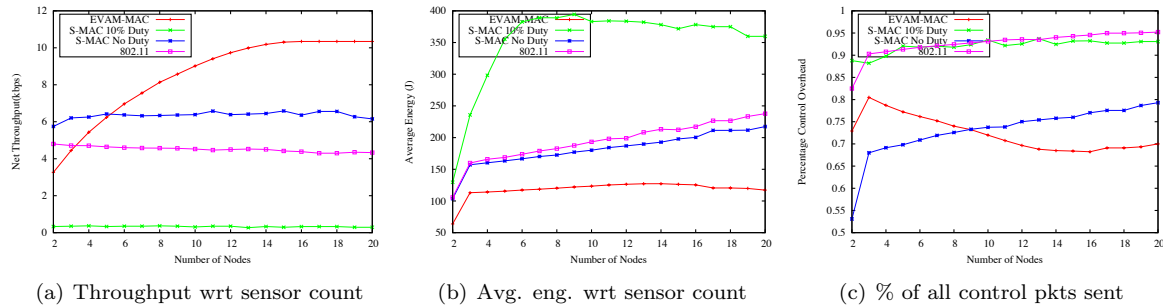


Figure 5: One-hop simulations for nodes with CBR and UDP agents Traffic rate = 8 pps

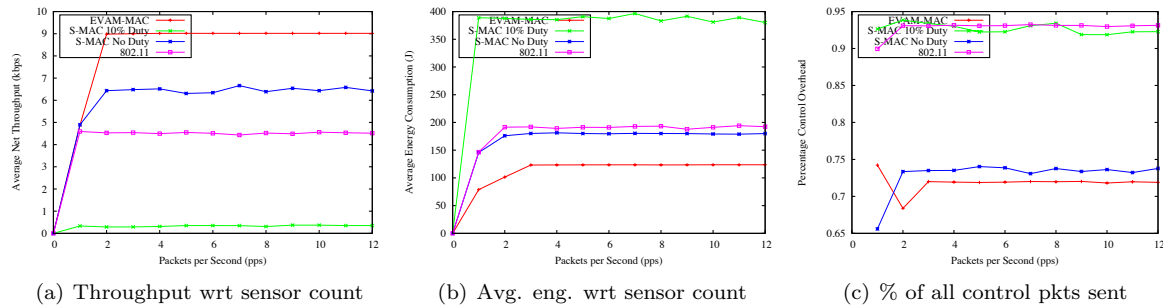


Figure 6: One-hop simulations for nodes with CBR and UDP agents Number of nodes = 10

rate) traffic generator. All nodes generate traffic at approximately similar time and unicast their data directly to the sink. Two simulation experiments were performed. In the first experiment, the sensor count is increased from 2 to 20 while the data traffic is kept constant to a value of 8 packets per second. As shown in Fig. 5a, the net throughput is compared for different MAC protocols as a function of sensor count. EVAM-MAC produces an exponential growth from sensor count 2 to 15 while it saturates around 10kbs. S-MAC was operated with a 10% duty cycle where its performance is the worst since the overhead of synchronization deteriorates the throughput severely. S-MAC with no duty cycle and 802.11 achieve similar performance. EVAM-MAC achieves higher throughput than all other protocols when the sensor count increase above 5. This is because the RTS sub-round timer is fixed to a value taking into account that 20 nodes may have data to send at the same time. When 5 nodes or less are transmitting, the RTS sub-round timer is large enough to lowers the throughput. However, EVAM-MAC is tailored for high sensor count and traffic application. Moreover, as shown in Fig. 5b, EVAM-MAC has the lowest energy consumed with respect to all MAC protocols. Fig. 5c, shows the average percentage of control packets sent by each MAC protocol of all nodes with respect to all packets

sent. EVAM-MAC has the lowest control overhead as sensor count increases above 9 nodes. This is due to the fact that the percentage of control bytes sent stays constant since the RTS sub-round timer is fixed while the number of DATA bytes increases. In the second experiment we fixed the number of nodes to 10 and we varied the number of packets sent per seconds from 1pps to 12pps. This experiment is performed to evaluate EVAM-MAC from low to high traffic conditions. As shown in Fig. 6a, the net throughput of EVAM-MAC is superior to all other protocol when the traffic generated by each node is equal or greater to 2pps. EVAM-MAC is also superior to all protocols compared in terms of energy consumed as shown in Fig. 6b, and percentage of control overhead sent shown in Fig. 6c. This proves the effectiveness of EVAM-MAC in both high sensor count and high traffic application in single hop networks.

The second application is a simulated event based system where nodes detecting an event are required to broadcast their Data packets to all nodes. After each node receives a data packet from every other node detecting the event it will unicast a Data packet to the sink. However, the number of nodes detecting the event is required to be known so that nodes receiving the broadcast information know when to unicast its Data packet. For EVAM-MAC, the number of nodes detecting the event can be known from the collected

RTS packets. However, for S-MAC protocol, nodes will have no a priori information about the number of nodes detecting the event. In this scenario, we assume a fixed number of nodes that is supplied to them at deployment. In real deployment, S-MAC will have to exchange some packets to know the number of nodes detecting the event. This will incur an overhead in energy consumption and will increase latency. EVAM-MAC uses the two round programmable feature that was discussed in Section 3.4. As shown in Fig. 6a, the net throughput is compared as the number of nodes detecting the event increases. EVAM-MAC throughput growth is exponential which reaches a value of 12kbs. EVAM-MAC's performance is superior to S-MAC when the sensor count increase above 5. Furthermore, as shown in Fig. 6b and Fig. 6c, EVAM-MAC has similar energy consumption values and control packet overhead to that of SMAC with 10% duty cycle. Thus, EVAM-MAC can achieve nearly double the throughput of S-MAC with no duty cycle with the same value of energy spent and control packet sent of S-MAC with 10% duty cycle.

4.2 Two-hop Benchmarks

To prove the correctness of EVAM-MAC for two hop scenarios, we simulated the proposed algorithm for scenarios 1 and 2 shown in Fig. 4. In both scenarios, detected event forms two clusters simultaneously where the maximum number of nodes in each cluster is equal to 12. In scenario 1, one node is in communication range of both clusters. In scenario 2, three nodes are in the interference range of both clusters. Each cluster has a sink that is in communication reach of all nodes belonging to that cluster. When an event is detecting all nodes will unicast their Data packets to the sink of the corresponding cluster. As shown in Fig. 7a, d, the net throughput is compared as the number of sensors detecting the event increases at both clusters. All protocols show a linear increase as the number of sensors detecting the event increases. However, EVAM-MAC presents superiority when the sensor count is around 16, which translates to 8 sensor nodes per cluster. The reason behind this is that EVAM-MAC shuts down the whole cluster while the other cluster is delivering data. As the number of sensor nodes in each cluster reaches 8, it becomes worthwhile to shut down the cluster as the contention produced from two neighboring clusters introduces latency. In Fig. 7b, e, the average energy consumed of EVAM-MAC has close performance with respect to S-MAC with no duty cycle. 802.11 has the worst energy consumed with respect to all compared protocols. Moreover, EVAM-MAC has greater control

packet overhead with respect to S-MAC (no duty cycle) as shown in Fig. 7c, where 802.11 has also the worst performance. However, this overhead is compensated for by the reduced number of dropped data packet which is evident in Fig. 7b,e.

4.3 Multi-hop Benchmarks

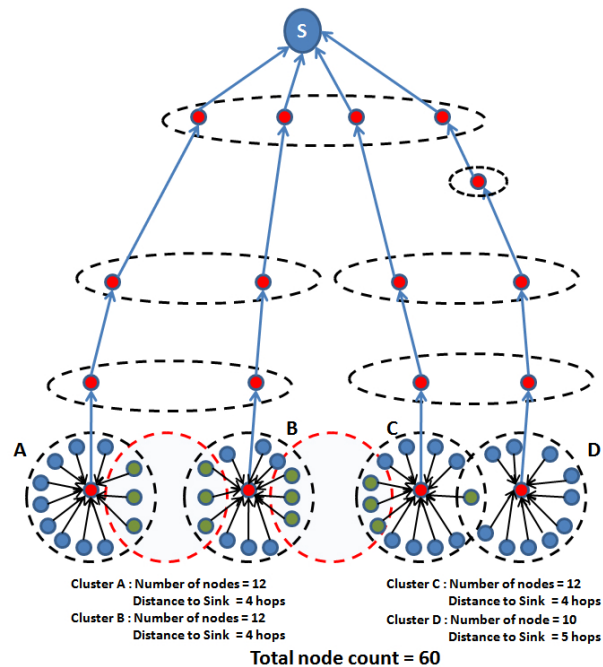


Figure 8: Multihop Layout

To evaluate EVAM-MAC in a multi-hop scenario, the proposed algorithm was simulated in a 60 node network as shown in Fig. 8. The maximum number of nodes detecting the event simultaneously is equal to 50. Four clusters exist as shown in Fig. 8 where cluster A, B and C are 4 hops away from the sink and each contain 12 nodes. Cluster D is 5 hops away from the sink and contains 10 nodes. To eliminate the overhead of routing, fixed routing tables were supplied to the nodes. Each node detecting the event will unicast its data to the local sink. Upon the reception of all data packet for all nodes detecting the event, the local sink will forward the aggregated Data packet through several hops until it reaches the final sink as shown in Fig. 8. Four metrics were used for this simulation: net throughput, average energy consumed, average control packet overhead and average latency. The average latency is the average time that all nodes take to transmit their packet to the final sink. As shown in Fig. 9a, we evaluated EVAM-MAC's throughput as a function of the number of

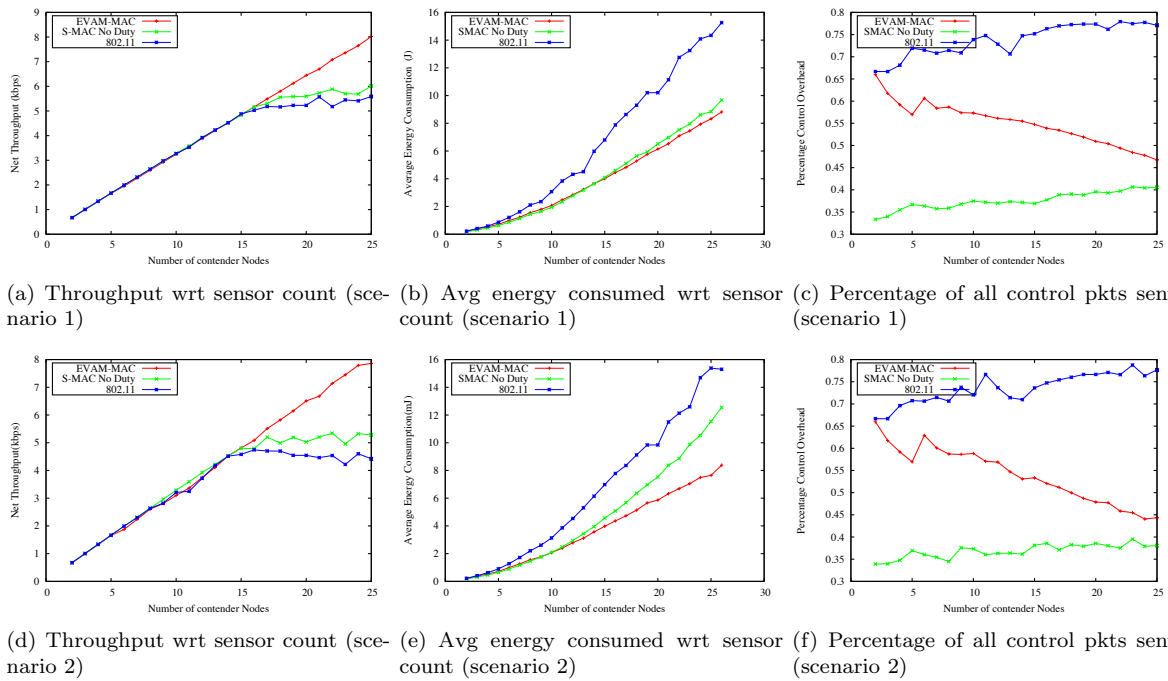


Figure 7: Two hop simulations for an event based system

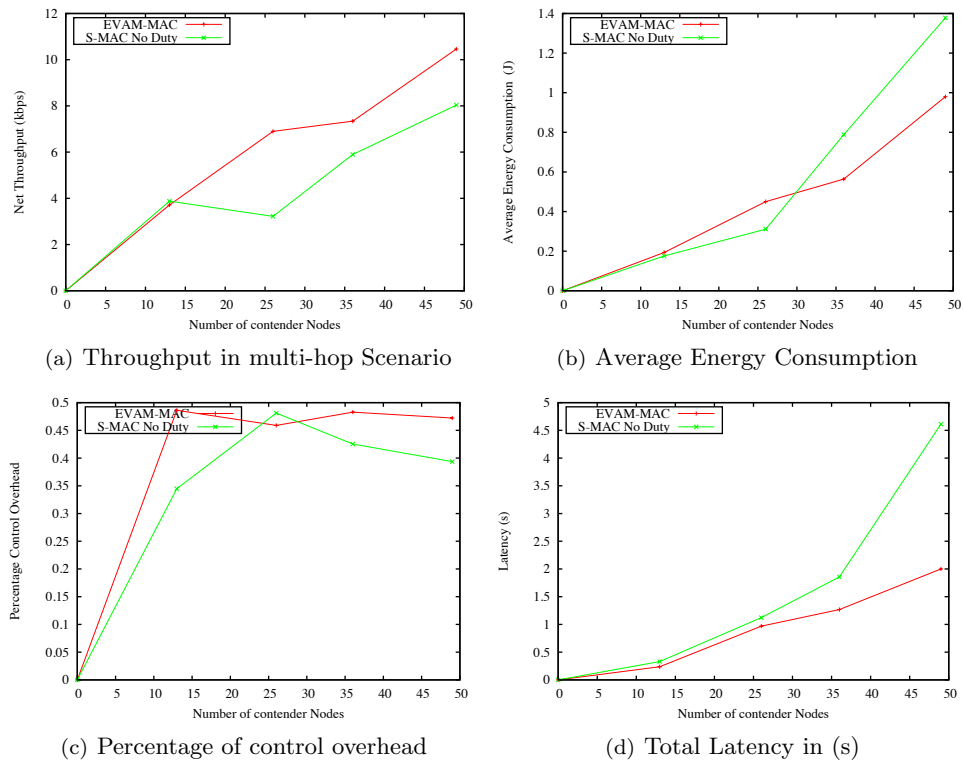


Figure 9: Multihop Simulation

sensors detecting the event. EVAM-MAC achieves higher throughput than S-MAC with no duty cycle as the number of sensors detecting the event grows above 15. We didn't compare our protocols with S-MAC with duty cycle since it is not scalable to this amount of nodes. Fig. 9b presents the average energy consumed by all nodes as sensor count increases. Initially, EVAM-MAC has similar performance as that of S-MAC with no duty cycle. However, when the sensor count increases above 25 EVAM-MAC has lower energy consumption. On the other hand, the control packet overhead of EVAM-MAC is higher than that of S-MAC as shown in Fig. 9c. However, this overhead didn't reflect negatively on energy consumption since the percentage of dropped packets is greatly reduced with EVAM-MAC. Moreover, EVAM-MAC has lower latency than S-MAC as shown in Fig. 9d. When the number of sensors detecting the event reaches 50, EVAM-MAC has a 3 to 1 advantage over S-MAC in latency. Thus, EVAM-MAC achieves lower latency and higher throughput without introducing a big overhead in energy consumption and control packets sent. EVAM-MAC is very desirable in event based systems as it relieves contention from multiple nodes contending for the Data packet delivery.

5 Conclusion and Future Work

In this work a novel MAC protocol tailored for event based systems for wireless sensor networks is presented. EVAM-MAC creates a contention free schedule for the collected measurements for transmission in a dynamic decentralized fashion. EVAM-MAC does not require any global synchronization or setup phase at deployment. This has the advantage of eliminating the overhead of nodes joining and leaving the network. EVAM-MAC shifts the contention generated by multiple sensor nodes detecting an event from the Data phase to the control phase where packet sizes are much smaller in size. Furthermore, EVAM-MAC provides a rich platform of programmable features that the application designers can utilize to achieve high performance operations. EVAM-MAC presented its superiority in throughput and energy conservation for single and multi hop scenarios when compared with S-MAC and 802.11.

Our future work includes performing a real wireless sensor implementation on wireless sensor nodes. EVAM-MAC will be tested on target localization application for wireless sensor networks. Furthermore, more simulations and experimentation with different topologies and higher sensor count is being performed.

References

- [1] P. Suriyachai, U. Roedig, A. Scott, "A Survey of MAC Protocols for Mission-Critical Applications in Wireless Sensor Networks", In *IEEE Communications Surveys & Tutorials*, vol.14, no.2, pp.240-264, Second Quarter 2012, doi: 10.1109/SURV.2011.020211.00036
- [2] W. Ye, J. Heidemann, and D. Estrin, "An Energy-Efficient Mac Protocol for Wireless Sensor Networks", *Proc. IEEE Conf. Computer and Communications Societies (INFOCOM 2002)*, vol. 3, pp. 1567-1567, Jun. 2002, doi: 10.1109/INFCOM.2002.1019408.
- [3] T.V. Dam and K. Langendoen, "An Adaptive Energy-Efficient Mac Protocol for Wireless Sensor Networks", *Proc. ACM Conf. Embedded Networked Sensor Systems (SENSYS '03)*, pp. 171-180, Nov. 2003, doi: 10.1145/958491.958512.
- [4] J. Polastre, J. Hill, D. Culler, "Versatile Low Power Media Access for Wireless Sensor Networks", *Proc. ACM Conf. Embedded Networked Sensor Systems (SENSYS '04)*, pp. 95-107, Nov. 2004, doi: 10.1145/1031495.1031508.
- [5] W. Ye, J. Heidemann, and D. Estrin. Medium access control with coordinated adaptive sleeping for wireless sensor networks. *IEEE/ACM Trans. Netw.*, 12(3):493-506, 2004. doi: 10.1109/TNET.2004.828953
- [6] Shams ur Rahman, Mushtaq Ahmad, Shafaat A. Bazaz, "A new Energy-Efficient TDMA-based MAC Protocol for Periodic Sensing Applications in Wireless Sensor Networks", In *IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 4, No 1, July 2012.
- [7] I. Rhee et al., "Z-MAC: a hybrid MAC for wireless sensor networks", *Proc. ACM Conf. Embedded Networked Sensor Systems (SENSYS '05)*, pp. 90-101, Nov. 2005, doi: 1098918.1098929.
- [8] Z. Merhi, M. Elgamel, and M. Bayoumi, "EVA-MAC: An Event based Medium Access Control for Wireless Sensor Networks", *Proc. of the IEEE Workshop on Signal Processing Systems (SiPs2011)*, pp. 61-66, Oct. 2011, doi:10.1109/SiPS.2011.6088950
- [9] Z. Merhi, M. Elgamel, and M. Bayoumi, "A Lightweight Collaborative Fault Tolerant Target Localization System for Wireless Sensor Net-

- works," IEEE Transactions on Mobile Computing, 29 Apr. 2009. IEEE computer Society Digital Library. IEEE Computer Society, doi: 10.1109/TMC.2009.81
- [10] The Network Simulator NS-2, <http://www.isi.edu/nsnam/ns/>, 2009.
- [11] LAN MAN Standards Committee of the IEEE Computer Society, "Wireless LAN medium access control (MAC) and physical layer (PHY) specification", IEEE, New York, NY, USA, IEEE Std 802.11-1997 edition, 1997.
- [12] G. Ahn et al., "Funneling-MAC: a localized, sink-oriented MAC for boosting fidelity in sensor networks", Proc. ACM Conf. Embedded Networked Sensor Systems (SENSYS '06), pp. 293-306, Nov. 2006, doi: 10.1145/1182807.1182837.
- [13] I. Rhee, A. Warriar and L. Xu, "Randomized dining philosophers to TDMA scheduling in wireless sensor networks" Technical report, Computer Science Department, North Carolina State University, Raleigh, NC, 2004.
- [14] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks", IEEE Trans. Wireless Communication, vol. 1, pp. 660670, Oct. 2002, doi:10.1109/TWC.2002.804190.
- [15] Chunyao FU et. al, "An Energy Balanced Algorithm of LEACH Protocol in WSN" in IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 1, January 2013.
- [16] J. Polastre, J. Hill, D. Culler, "Versatile Low Power Media Access for Wireless Sensor Networks", Proc. ACM Conf. Embedded Networked Sensor Systems (SENSYS '04), pp. 95-107, Nov. 2004, doi: 10.1145/1031495.1031508.
- [17] M. Buettner et al., "X-MAC: A Short Preamble MAC Protocol For Duty-Cycled Wireless Sensor Networks", In Proc. of 5th ACM Conference on Embedded Networked Sensor Systems (SenSys 2006), November 2006.
- [18] W. Ye, F. Silva, and J. Heidemann, "Ultra-Low Duty Cycle MAC with Scheduled Channel Polling" In Proc. of 5th ACM Conference on Embedded Networked Sensor Systems (SenSys 2006), November 2006.
- [19] Yanjun Sun, Omer Gurewitz, and David B. Johnson. "RI-MAC: A receiver initiated asynchronous duty cycle MAC protocol for dynamic traffic loads in wireless sensor networks". In Proc. of 6th ACM Conference on Embedded Networked Sensor Systems (SenSys'08), 2008.do:10.1145/1460412.1460414
- [20] Lei Tang, Yanjun Sun, O. Gurewitz, D.B. Johnson, "PW-MAC: An energy-efficient predictive-wakeup MAC protocol for wireless sensor networks," In Proc. of the 30th IEEE International Conference on Computer Communications ((INFOCOM 2011), pp.1305-1313, 10-15 April 2011, doi: 10.1109/INFCOM.2011.5934913
- [21] Beakcheol Janga, Jun Bum Limb, Mihail L. Sichitiuc, "An asynchronous scheduled MAC protocol for wireless sensor networks" In Elsevier Journal on Computer Networks, 2012

Zaher M. Merhi received the B.Sc. degree and the M.Sc. degree in computer engineering from the University of Balamand, Lebanon, in 2002 and 2005. He also received the M.Sc. and the PhD degree in computer engineering from the University of Louisiana at Lafayette, USA, in 2007 and 2010. Since 2010 he joined the Lebanese International University (LIU) as an assistant professor at the computer and communication engineering (CCE) department. Currently he is the associate chairperson of the CCE department at LIU. His research interests include low power embedded systems, distributed DSP applications for wireless sensor networks, data fusion techniques and medium access controls (MAC) for wireless sensor networks.

Mohamed A. Elgamel received the B.Sc. degree in computer science from Alexandria University, Alexandria, Egypt, in 1991, the M.Sc. degree in computer engineering from Arab Academy for Science and Technology (AAST), Alexandria, Egypt, in 1998, the M.Sc. and the PhD degrees in computer engineering from University of Louisiana at Lafayette, Louisiana, USA, in 2000 and 2003 respectively. He is currently an Associate Professor at the Faculty of Computing and Information Technology, AAST. In 2013, he was a visiting professor at The International Telematic University (UNINETTUNO), Rome, Italy. From 2003-2010, he has been a Graduate Faculty at The Center for Advanced Computer Studies, University of Louisiana at Lafayette, USA. From 1999 to 2003, he served as a Research Assistant and Adjunct Faculty at the Center for Advanced Computer Studies, University of Louisiana at Lafayette, USA. From 1991 to 1999, he was a Senior Analyst Programmer

at Arab Academy for Science and Technology, Egypt. His research interests include CAD tools, Embedded Systems, Wireless Sensors Networks, Software Engineering, and digital video processing. His research was supported by funds from federal agents like DOE, NSF and the Louisiana Governor ITI. Dr. Elgamel is a senior IEEE member. He is also a member on the IEEE Technical Committee on Communication and IEEE-TC on VLSI. He served on the program committees and chaired sessions of several conferences and as a reviewer/Reviewer Committee Member for several conferences and journals. Dr. Elgamel received the 2002 Richard E. Merwin Student Scholarship for demonstrating outstanding involvement in an IEEE-CS and excellence in academic achievement. He was the first place winner in the student paper contest in the 46th IEEE International Midwest Symposium on Circuits and Systems, (MWSCAS2003). He was the president of the IEEE Computer Society student chapter at the University of Louisiana at Lafayette.

Samih Abdul-Nabi received his B.E. degree in Electrical Engineering in 1990 from the Lebanese University, Beirut, Lebanon and his M.S. degree in Computer Sciences in 1993 from the University of Montreal, Canada. Since 1993, Mr Abdul-Nabi worked in many related engineering fields in both Canada and the U.S. His last industrial assignment was at Diebold Inc. as Senior Systems Engineer providing technical support to one of the largest banks in Canada. His is currently a lecturer at the Lebanese International University working on some research topics that includes Network Coding and encryption.

Amin A. Haj-Ali received his B.E. degree in Computer and Communications Engineering in 1993 from the American University of Beirut, Beirut, Lebanon, his M.S. degree in Electronics and Computer Control Systems in 1994, and his Ph.D. in Electrical Engineering (Systems and Soft Computing) and Computer Science (Minor) in 2002 from Wayne State University, Detroit, USA. Dr. Haj-Ali earned a Masters in IT Project Management from George Washington University, Washington DC, USA. Since 1996, Dr. Haj-Ali worked in many related engineering fields in both Lebanese and international corporations. His last industrial assignment was at DaimlerChrysler

Corporation Scientific Labs and Proving Grounds, Auburn Hills, MI USA where has registered a patent (US07043963B2). Currently Dr. Haj-Ali is the Associate Dean of the School of Engineering and Associate Professor at the Lebanese international University. His research interest includes Machine Vision, Embedded Systems, Fuzzy Systems and Neural Networks.

Magdy A. Bayoumi received the B.Sc. and M.Sc. degrees in electrical engineering from Cairo University, Cairo, Egypt, the M.Sc. degree in computer engineering from Washington University, St. Louis, and the Ph.D. degree in electrical engineering from the University of Windsor, Windsor, ON, Canada. He is the Director of the Center for Advanced Computer Studies (CACs) and Department Head of Computer Science Department, University of Louisiana, Lafayette. He is also the Z.L. "Zeke" Loflin Eminent Scholar Endowed Chair at the Center for Advanced Computer Studies, University of Louisiana, Lafayette, where he has been a faculty member since 1985. His research interests include VLSI design methods and architectures, low power circuits and systems, digital signal processing architectures, parallel algorithm design, computer arithmetic, image and video signal processing, neural networks, and wideband network architectures. Dr. Bayoumi was the vice president for the technical activities of the IEEE Circuits and Systems Society and the chairman of the Technical Committee on Circuits and Systems for Communication and the TC on Signal Processing Design and Implementation. He was a founding member of the VLSI Systems and Applications Technical Committee and was its chairman. He is a member of the Neural Network and the Multimedia Technology Technical Committees. He was an Associate Editor of the Circuits and Devices Magazine, the IEEE Transactions on Very Large Scale Integration (VLSI) Systems, the IEEE Transaction on Neural Networks, and the IEEE Transaction on Circuit and Systems II: Analog and Digital Signal Processing and Integration. He was also the chairman of many conferences and workshops including MWSCAS'94, GLSVLSI'98, MWSCAS'03, ISCAS'07, and ICIP'09.

A Critical Review of Development of Intelligent Tutoring Systems: Retrospect, Present and Prospect

Dr. Neelu Jyothi Ahuja¹ and Roohi Sille²

¹University of Petroleum and Energy Studies, Dehradun.
Assistant Professor, Deptt of CSE

²University of Petroleum and Energy Studies, Dehradun.
MTech(AI & ANN), Deptt of CSE

ABSTRACT:

This paper introduces, Intelligent Tutoring Systems along with their typical architecture, developmental history, past and present systems and concludes with a broad discussion on wide-spanning focus areas for future developmental research. A critical analysis of the developmental history highlighting the theme behind the developed systems, their purpose and the key ITS concept, have been presented. A closer look revealed that, development of a certain concept proved to become a turning point for all future developments of that era. All such key concepts and subsequent developments have been examined. The paper provides recommendations and pointers, to the areas that need to be probed further and drilled down to establish ITS success for generations to come.

KEY WORDS:

Intelligent Tutoring Systems (ITS), Computer Assisted Instruction(CAI), Case Based reasoning(CBR), Affective Tutoring Systems(ATS), Expert Systems, Expert Systems Development.

1. INTRODUCTION

The amalgamation of Artificial Intelligence techniques into education, producing educationally useful

computer artifacts dates back to early 1970s. Over a large spectrum of incremental developments, they have taken various forms, one amongst which is Intelligent Tutoring Systems (ITSs). It is a computer-based program not only to emulate a ‘human tutor’, but to personalize the instructions based on the background and progress of each individual learner. There has been evolution from a very primitive form of computer-assisted instruction, ranging through various forms of e-learning systems, progressing to form learner adaptive systems, to modern day ITS, with significant development in their user interface as well, highlighting and facilitating a smooth cognitive interaction of man and machine. The systems have transformed to become a true intersection of computer science, cognitive psychology and educational research. They have offered various focus points in system development across various time periods highlighting research areas, on and off being addressed by researchers from time to time.

1.1. Architecture of a Typical ITS System

A typical ITS, has the following four basic components [1]. The section below lists them with their functionality, individually and then by way of their integration.

1.1.1. The Domain model

1.1.2. The Student model

1.1.3. The Tutoring model, and

1.1.4. The User interface model

1.1.1. Domain Model: The domain model (also known as the cognitive model/expert knowledge model) consists of the concepts, facts, rules, and problem-solving strategies of the domain in context. It serves as a source of expert knowledge, a standard for evaluation of the student's performance and diagnosis of errors.

1.1.2. Student Model: The student model is an overlay on the domain model. It emphasizes cognitive and affective states of the student in relation to their evolution as the learning process advances. As the student works step-by-step through their problem solving process, the system engages itself in model tracing process. Anytime there is any deviation from the predefined model, the system flags it as an error.

1.1.3. Tutoring Model: The tutor model (also called teaching strategy or pedagogic module) accepts information from the domain and student models and devices tutoring strategies with actions. This model regulates instructional interactions with student. It is closely linked to the student model, makes use of knowledge about the student and its own tutorial goal structure, to devise the pedagogic activity to be presented. It tracks the learner's progress, builds a profile of strengths and weaknesses relative to the production rules (termed as 'knowledge-tracing').

1.1.4. User Interface Model: This is the interacting front-end of the ITS. It integrates all types of information needed to interact with learner, through graphics, text, multi-media, key-board, mouse-driven menus, etc. [2]. Prime factors for user-acceptance are user-friendliness and presentation. The Figure 1 presents a typical ITS architecture.

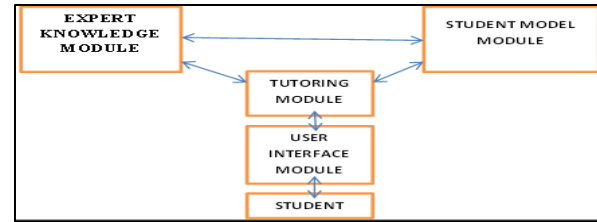


FIG 1: Typical Architecture of an ITS

2. CHRONOLOGICAL REVIEW OF ITS DEVELOPMENT

This section of the paper presents ITS development across past (1970s-1999) and present (2000-2013).

In the past decade, there has been tremendous growth in the field of expert systems and ITS with student modeling as a research area maturing sufficiently constituting a very promising technology for personalization and adaptivity of e-learning systems. Since 1960's to present, ITS have been heralded as one of the most promising approaches to deliver individualized instructions. In the early 1960, programmed instruction, enhancing learning for low aptitude individuals, was educationally fashionable, moving towards structured and goal oriented instruction [1].

Dawn of 1970's saw a new era of ITS development with knowledge representation, student modeling, Socratic tutoring, skills and strategic knowledge, buggy library, expert systems and genetic graph. "Bug Library" is a collection of mistakes. In genetic graph, "Genetic" related to the notion of knowledge being evolutionary, and graph denoted the relationships between parts of knowledge expressed as links in a network. In 1980's, the emphasis in ITS development was case-based reasoning, more buggy based systems, discovery worlds, progression of mental models, simulation, natural language processing, authoring systems and systems based on model tracing. Model tracing Tutors contained a cognitive model or

simulation of an expert's correct thinking in the domain [3].

In 1990's focus shifted to learning theory that embodied concepts such as learner control, collaborative as against individual learning, information processing and virtual reality as against situated learning. Unlike individual learning, people engaged in collaborative learning capitalizing on one another's resources and skills. Both novice and master are active participants in the learning environment.

In the present years technological resources have been integrated with education. However, the integration of educational technology at early childhood education is a more recent trend compared to at other levels of education. From year 2000 to 2013, important issues related to ITS development concentrated on student modeling approach, learning through games, adaptation to emotional state of user, web based tutoring systems, knowledge modeling by fuzzy linguistic information, WIMP interfaces, summary assessment techniques, motion capture technology, interrelation between person's cognitive load and pupil's size and education data mining.

The section below represents a retrospective developmental account of ITS between the period 1970 and 1999.

Basic Instructional Program (1970) employed teaching procedural skills in learning programming language BASIC. Exercises were dynamically and individually selected per user using Curriculum Information Network (CIN) [4]. **Carbonell's SCHOLAR (1970)** used semantic net to represent domain knowledge as well as the student model.[5]. Collins in 1975 outlined set of tutorial rules for **Socratic tutoring**. One such system was **WHY**. It stores domain knowledge in script hierarchy containing stereotypical sequences of events [6].

WEST [7] helped students to improve arithmetic expression manipulation skills. It was called issue-based tutoring. **SOPHIE (Sophisticated Instructional Environment)** assisted learners in developing electronic troubleshooting skills. SOPHIE I, SOPHIE II, SOPHIE III have extended the environment of their predecessors. [8]. **BUGGY (1978)** employed buggy library approach for diagnosis of student errors (bugs). It was a framework for modeling misconceptions underlying procedural errors in addition and subtraction exercises offered to student for solving. [8]. **DEBUGGY [9]** was an offline version of a system based on BUGGY using the pattern of error. IDEBUGGY developed by Burton in 1982 was an on line version to diagnose student's procedure bit by bit while giving the learner a new problem to solve at each step. Limitation of buggy library was its inability to anticipate all possible misconceptions. **MYCIN [10]** was a rule-based expert system for diagnosing certain Infectious diseases such as meningitis. Using the learning of MYCIN, **GUIDON** was constructed by Clancey in 1979 to interface with MYCIN for tutoring, interactively presenting the rules in the knowledge base to a student [11]. **WUSOR** was the name of the on-line coach for the game WUMPUS, developed by Stansfield, Carr and Goldstein in 1976 [12]. LISP Tutor by Anderson Boyle and Reiser and a Geometry Tutor by Anderson Boyle and Yost arrived in mid-1980 employed the approach of model tracing. [13-14]. **PROUST** by Johnson and Littman Soloway in 1984 diagnosed non-syntactic student errors in PASCAL [15]. **PIXIE** developed by Sleeman in 1987 is an online ITS based on Leeds Modeling System (LMS) having a diagnostic model for determining sources of errors in algebra due to incorrect (mal) rules that are inferred from basic principles and bugs at abstraction level [16.]

In late 1980 arrived the **Case-based Reasoning (CBR)** research by Schank and Kolodner which had a more adaptive learning environment, with the advantage of being suitable to domains where there are too many ways in which the rule can be applied (e.g., programming, game playing) and suggests approximate answers to complex problems. [17-18].

The year 1990 brought the new trend of graphic simulations. Hauk Mack III was a system that expanded number of components and complexity of animations by orders of magnitude [19].

The other areas of research and development that gained prominence were Natural Language Processing (NLP) and authoring shells. SOPHIE was built on a powerful and original NLP technique developed by Richard Burton; called Semantic Grammar. It represented a powerful combination of carefully selected keywords with algorithms that searched the context for meaningful variables and objects. Authoring shells are kind of e-learning systems that feature authoring environments for system users, simplify the software development life cycle. Domain knowledge in such systems can be represented by using different knowledge representation specifications. [9][20][21].

3. THE PRESENT SCENARIO OF ITS DEVELOPMENT

In recent years, progress has been towards providing adaptivity and personalization in computer based education through student modeling, mobile technologies, educational games and standalone educational applications.

An adaptive educational system has to provide personalization to the specific needs, knowledge and background of each individual student which is challenging since students not only have different learning needs, but also different learning

characteristics. The section below lists each of these concepts with their applicability in present day ITS development. The major approaches introduced were overlay, perturbation, stereotypes, machine learning techniques, cognitive theories, constraint based model, fuzzy student modeling, Bayesian network, and ontology student modeling [22]. Few of these have been briefly stated below. The overlay model was invented by Stansfield, Carr, and Goldstein in 1976 and has been used in many systems ever since [23]. The main assumption underlying it is that a student may have incomplete but correct knowledge of the domain. Overlay models are inadequate for sophisticated models because they do not take into account the way users make inferences, or integrate new and old knowledge, and change representational structures with learning. Stereotypes were introduced to user modeling by Rich in 1979 in the system called GRUNDY [24]. The main idea of stereotyping was to cluster all possible users of an adaptive system into several groups according to certain characteristics that they typically shared [25]. A perturbation student model is an extension of the overlay model that represents the student's knowledge as including possible misconceptions as well as a subset of the expert's knowledge [26]. It represents learners as the subset of expert's knowledge, like the overlay model, plus their mal-knowledge [27]. This extension allows for better remediation of student mistakes, since the fact that a student believes something that is incorrect is pedagogically significant [28]. The processes of observation of student's action and behavior in an adaptive and/or personalized tutoring system, and of induction, should be made automated by the system. A solution for this is machine learning, which is concerned with the formation of models from observations and has been extensively studied for

automated induction. [29]. The cognitive theory attempts to explain human behavior during the learning process by understanding human's processes of thinking and understanding [57]. The Constraint-Based Model (CBM) proposed by Ohlsson in 1996 is based on Ohlsson's theory of learning from errors, and proposes that a learner often makes mistakes when performing a task, even when he/she has been taught the correct way to do it. Fuzzy logic is able to handle uncertainty in everyday problems caused by imprecise and incomplete data as well as human subjectivity. Fuzzy Student Modeling was applied, by Stathacopoulou et al. in 2005 to a discovery-learning environment that aimed to help students to construct the concepts of vectors in physics and mathematics [30]. Several student models have been built based on ontologies. These support the representation of abstract concepts and properties so as to be easily reused and, if necessary, extended in different application contexts [31]. A glimpse of few significant systems of the era is presented below:

Adaptive Intelligent Web Based Education Systems (AIWBES) were developed as an alternative to traditional e-learning environments according to 'onesize-fits-all' approach [32][58]. **Affective tutoring systems (ATS)** [33]. The system utilizes a network of computer systems, prominently, embedded devices to detect student emotion and other significant bio-signals and adapt to the student's mood and display emotion via a life-like agent called Eve, whose tutoring adaptations are guided by a case-based method for adapting to student states - confused, frustrated or angry [34]. **Multi Criteria decision model** has been employed to integrate expert's knowledge modeled by fuzzy linguistic information, enhancing accuracy of diagnosis for adaptation of computerized test of the student competence level.

Pen-based tutoring systems are based on WIMP (windows, icons, menu & pointer) interfaces. Newton's Pen is a "statics tutor" implemented on a "pen top computer," a writing instrument with an integrated digitizer and embedded processor. This project entailed the development of sketch understanding techniques and user interface principles for creating pedagogically sound instructional tools for pen top computers. Development on the pen top platform presented novel challenges because of limited memory and computational power resources [34].

Automatic Summary Assessment has been a widely used mechanism. Several techniques such as latent semantic analysis (LSA), n-gram co-occurrence and BLEU have been proposed to support automatic evaluation of summaries [35]. Landauer et al in 1998 first developed latent semantic analysis (LSA) in the late '80s with the purpose of indexing documents and information retrieval [36]. LSA works by using a matrix to capture words and frequency of the words appearing in a context that is transformed using Singular Value Decomposition (SVD). Based on the result of Landauer's experiment, LSA is capable of producing acceptable results. However, LSA does not make use of word order as Landauer claims that word order is not the most important factor in collecting the sense of a passage. Pérez et al. in 2004 modified the BLEU algorithm, which was originally developed for ranking machine translation systems, into one that is capable of marking students' essay. [37]. Lin and Hovy in 2003 conducted a study on using the two machine translation evaluation techniques, BLEU and NIST's n-gram co-occurrence scoring procedures, on the evaluation of summaries to measure the closeness of the candidate to the reference summary [38]. With the recent success of e-learning and advances in other areas such as Information Extraction (IE) and NLP,

automatic assessment of summary writings has become possible.

Handwriting Based Intelligent Tutors use handwriting input offering several affordances for students that traditional typing-based interactions do not [39].

Educational Data Mining (EDM) is concerned with developing, researching, and applying computerized methods to detect student access patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data within which they exist [40].

Motion Capture Technology is being used in automated lesson generation systems for example one such system is 'Dance Learning from Bottom-Up Structure (DL-BUS)' for guiding beginners to learn basic dance movement, analyzing the dance to generate a two-phase lesson (phase-1 to divide dance into small segments and phase -2 to combine patterns in temporal order) providing suitable cognitive load thus offering an efficient learning experience.[41].

A level ahead is an **Intelligent Pupil Eye Analysis System**, involving the interrelation between person's cognitive load and pupil size. This sensitivity of the pupil can provide exhaustive data about the cognitive loads. Different works such as by Klingner et al., in 2008; Partala and Surakka, in 2003; Valverde et al., in 2010; Klingner, in 2010; Just and Carpenter, in 1993; Backs and Walrath, in 1992; and Porter et al., in 2007 demonstrate that task-induced dilations can serve as reliable proxies for cognitive load, and the sizes of blink pupil dilations reliably reflect a diverse scale of the difficulty of different activities thus validating pupillary dilations. [42-49].

Non-crisp learner responses that are uncertain usually belong to completely understanding or not understanding case for the content of learned

courseware. One of the Response Theory was Personalized Learning Item Response Theory (PEL-IRT), which including the fuzzy aspects, transformed into **Fuzzy Item Response Theory (FIRT)**, proposed by Chih-Ming Chen and Ling-Jiun Duh correctly estimated learner ability via the fuzzy inference mechanism [50-51].

UZWEBMAT: (Turkish abbreviation of Adaptive and Intelligent WEB based Mathematics teaching-learning system) -teaches secondary school level permutation, combination, binomial expansion and probability. [52]

4. FUTURE ASPECTS AND RECOMMENDATION:

The section outlines the areas where considerable amount of development is needed and awaited to make ITS systems live up to their objective. The expectations raised by these systems during initial days have made it more essential for a fool-proof system to evolve in the current century. A few such areas needing serious thought and contemplation by engineers/ research scientists are stated below.

Computer Assisted Assessment is a long standing problem that has attracted interest from research community since sixties and has not fully been resolved yet. The **ASSISTment** software provided hints and scaffolding in response to students' problem solving errors. The students performed better on a post-test than peers who completed their homework in traditional paper-and-pencil form, meaning that they did not receive immediate feedback and assistance on the problems. The results were encouraging, although limited by the relatively brief nature of the intervention. Previously, many researchers put efforts into e-learning systems with personalized learning mechanism to aid on-line learning. However, most systems focused on using learner's behavior, interests,

and habits to provide personalized e-learning services, but neglected the match between learner's ability and the difficulty level of the recommended courseware. Frequently, unsuitable courseware caused learner's cognitive overload or disorientation during learning. This area needs significant intervention by research community.

Adaptive abilities of ITS, are still not high enough, particularly regarding modes of practical problem-solving and support to learner in this process. Enriched adaptation techniques are required with focus on student's behavior. From the early years of systematic use of instructional design, educational scientists desired to use the results of artificial intelligence to support authors, developers and researchers, in their pedagogical work to create 'automatic' course designing machines. The objective was to make the built-in process more and more responsive and adaptive to the tuition circumstances, resulting in the design of a more intelligent training material. The last thirty year development in this discipline is still in an emerging phase.

The evaluation of ITS is an important though often neglected development stage. There are many evaluation methods available but literature does not provide clear guidelines for the selection of evaluation method(s) to be used in a particular context. Conventional computer programs are sometimes verified and validated through formal proofs of correctness. However, this technique is unsuitable for AI programs which deal with analytically intractable problems, represented as incompletely specified functions.

Extensively validated research in cognition, perception, and learning as indicated by Jay, in 1983; Jonassen & Hannum, in 1987; Larsen, in 1985 suggests ways to design and improve educational

programs, particularly the interface and user-related features. The area of human computer interaction holds potential that needs to be explored for channelized and focused development [53-55].

An expert's knowledge/inspection (called evaluation) is used as an explicit standard for judging a program. Due to ITS complex and dynamic behavior it is not as easy. Moreover there are known bottlenecks in extracting expert knowledge such as limited number of experts, their varied degrees of expression, difficulty in converting the knowledge gained from experience into a documented form etc. This is an area to further introspect.

The use of machine learning techniques can greatly improve the dynamic construction and updating of student models. There is hardly any research done for developing ITS for blind students, although there are systems developed for hearing impaired participants [56]. There are other set of disabilities, where there is a need for ITS to train. Training/exploring potential in creative art in creative arts is an area which has limited intervention as of now.

The main limitation of model tracing with respect to ill-defined domains is that, for some domains there are no clear strategies for finding solutions and it can therefore be difficult to define an explicit task model. Moreover, for complex domains, one would need to determine a large number of rules and solution paths, and designing a set of rules or a state space for each task would be very time consuming.

ITS does not appreciate or fails to offer encouragement at the time of student need. It fails to provide help in context of confusion causing to perceive negative emotions for her/his own actions. The shift has been made towards the methodologies instead of the much needed attention towards the student or the domain. Social learning/collaborative

learning is an area where a number of users come-together, collaborate, discuss and enhance their degree of understanding about a topic. There is a need for ITS researchers to explore the possibility and develop a framework for integrating social networking agents. There is a need to drill down and analyze emotional states of the learner and accordingly align the focus as well as learning/teaching strategies of ITS. There is immense future research direction embedded in it. With an increased demand for portable devices, various hand-held intelligent tutoring systems promise rich dividend.

5. CONCLUSION

The paper chronologically presents the development of ITS. It presents the retrospect present and the prospect of ITS. Over the period, they have gradually moved closer to the individual student learning need. Adaptability and user-friendliness have been the key concepts. Further, the prospective areas for future ITS development have been outlined for recommended research work. Human computer interaction has emerged as an area offering definite potential and demanding intervention by research scientists.

ACKNOWLEDGEMENT

The authors are grateful to the management of the University of Petroleum and Energy Studies, Dehradun for support in preparation of this article, and for granting permission to publish it.

REFERENCES

- [1]. Hyacinth S.N, 1990, Intelligent tutoring Systems: An Overview, AI Review, pp 251-277.
- [2]. Padayachee I. 2002. Intelligent Tutoring Systems: Architecture and Characteristics.
- [3]. Shute, V. J., Psotka, J., 1994., Intelligent Tutoring Systems: Past, Present, and Future,

Human resources directorate manpower and personnel research div., pp. 2-52

[4] Barr A., Beard M., Atkinson R.C., 1974, A Rational and Description of the BASIC Instructional Program, Tech Report, Stanford.

[5] Carbonell, J.R, 1970, AI in CAI: An AI Approach to CAI, IEEE Transactions on Man-Machine Systems, V. 11, pp, 190

[6] Collins, A. M., & Loftus, E. F.,1975, A Spreading Activation Theory of Semantic Processing. Psychological Review, 82, 407-428.

[7]. Brown, J. S., Burton, R. R., 1975. Multiple representations of knowledge for tutorial reasoning, Representation and understanding: Studies in Cognitive Science, pp.311-349.

[8]. Brown, I. S., Burton, R. R., Dekleer, J., 1982, Pedagogical natural language, and knowledge engineering techniques in SOPHIE I, II, and m, Intelligent Tutoring Systems, pp. 227-282.

[9]. Burton, R. R., 1982, Diagnosing bugs in a simple procedural skill, Cognitive Science, 4, pp. 157-183.

[10]. Shortliffe, E. H., 1976, Computer based medical consultations, MYCIN.

[11] William J. Clancey, 1979, Dialogue Management for Rule-Based Tutorials, Proc: International Joint Conference on AI, pp 155-161.

[12]. Stansfield, J. C., Car, B., Goldstein, L P., 1976, Wumpus advisor I: A first implementation of a program that tutors logical and probabilistic masoning skills, AI Lab Memo 381, MIT, Cambridge, Massachusetts.

[13] Anderson, J. R., & Reiser, B. J.,1985, The LISP tutor, Byte, 10, pp 159-175.

[14] Anderson, J. R., Boyle, C. F., & Yost, G ,1985, The geometry tutor, Proc: International Joint Conference on AI, pp 1-7.

[15]. Johnson, W. L., & Soloway, E. M., 1984, PROUST: Knowledge-based program debugging,

Proc: The 7th International Software Engineering Conference, Orlando, Florida, pp. 369-380.

[16]. Sleeman, D. H., 1987, PIXIE: a shell for developing intelligent tutoring systems, *AI & education: Learning environments and intelligent tutoring systems*, pp. 239-265.

[17]. Schank, R.C., 1982, *Dynamic Memory: A Theory of Learning in Computers and People*, Cambridge University Press, New York.

[18]. Kolodner, J. L. , 1988, Proc: 1st Case-Based Reasoning Workshop, Los Altos, CA.

[19]. Kurland et al., 1989, "HAWK MACHIII Explanations of the Receiver, Troubleshooting Tree, Technical Report, BBN Systems and Technologies, Cambridge, Massachusetts.

[20]. Itko B., et al., 2009, Dynamic test generation over ontology-based knowledge representation in authoring shell, *Expert Systems with Applications*, Vol 36, No. 4, pp 8185-8196.

[21]. Rosic, M, Vlado, G, Branko, Z., 2004, Intelligent authoring shell based on web services, Proc: INES -8th International conference on Intelligent Engineering Systems, pp 50-55.

[22]. Konstantina C, Virvou, M., 2013, Student modeling approaches: A literature review for the last decade, *Expert Systems With Applications*, Vol 40, no 11 pp. 4715-4729.

[23]. Stansfield, J., B.Carr, and I.P.Goldstein, 1976, Wumpus Advisor I: A first implementation of a Program that Tutors Logical and Probablistic Reasoning Skills, MIT AI Lab Memo No. 381.

[24]. Rich, E., 1979, User Modeling via Stereotypes, *Cognitive Science* 3, pp 329-354.

[25]. Tsiriga, V., & Virvou, M., 2002, Initializing the student model using stereotypes and machine learning. Proc: IEEE international conference on system, man and cybernetics pp. 404-409.

[26]. Mayo, M. J., 2001, Bayesian student modeling and decision-theoretic selection of tutorial actions in ITS, PhD Thesis.

[27]. Nguyen, L., Do, P., 2008, Learner model in adaptive learning, Proc: World academy of science, Engg and Technology, pp. 396-401.

[28]. Surjono, H., Maltby, J., 2003, Adaptive educational hypermedia based on multiple student characteristics, Proc: 2nd international conference on web based learning, Australia, pp. 442-449.

[29]. Webb, G., et al., 2001, Machine-learning for user-modeling, *User Modeling and User-Adapted Interaction*, 11, pp19-29.

[30]. R. Stathacopoulou, et al., 1999, Neural Network-based Fuzzy Modeling of the Student in Intelligent Tutoring Systems, Proc: Inter. Joint Conference on Neural Networks, Washington, IEEE Catalog No: 99CH36339C, 1999, paper #679.

[31]. Clemente, J., Ramírez, J., de Antonio, A., 2011, A proposal for student modeling based on ontologies and diagnosis rules, *Expert Systems with Applications*, 38(7), pp 8066-8078.

[32]. Brusilovsky, P., Peylo, C., 2003, Adaptive and intelligent web based educational systems, *International Journal of AI in Education*, 13, pp 156-169.

[33]. R. W. Picard, *Affective Computing*, MIT Press, 1997.

[34]. A. Sarrafzadeh., 2008, How do you know that I don't understand? A look at the future of intelligent tutoring systems, *Computers in Human Behavior*, Vol 24, no 4, pp 1342-1363.

[35]. He, Yulan et al., 2009, Automatic summary assessment for intelligent tutoring systems, *Computers and Education*, 53(3), pp. 890-899.

[36]. Landauer, T. K., Foltz, P. W., & Laham, D. 1998, Introduction to Latent Semantic Analysis, *Discourse Processes*, 25, pp 259-284.

[37].P'ere, D., 2005, Automatic Assessment of Students' free-text Answers underpinned by the Combination of a BLEU-inspired algorithm and Latent Semantic Analysis, *American Asso. for AI*, pp 101-119.

- [38] Lin, C.-Y. and E.H. Hovy. 2003, Automatic Evaluation of Summaries using n-gram Co-occurrence Statistics. Proc: HLT-NAACL conf, Edmonton, Canada.
- [39]. L. Anthonyn, J. Yang, K. R. Koedinger, 2012, A paradigm for handwriting-based intelligent tutors, *IJHCS*, Vol 70 no 11, pp 866-887.
- [40]. C. Romero, S. Ventura, 2012, Data mining in education, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol 3, no 1, pp 12-27.
- [41]. Y. Yang, H. Leung, L. Yue, L. Deng, 2013, Generating a two-phase lesson for guiding beginners to learn basic dance movements, *Computers and Education*, 61, pp 1-20.
- [42]. Klingner, J., Kumar, R., Hanrahan, P., 2008. Measuring the Task-Evoked Pupillary Response with a Remote Eye Tracker. Proc: The 2008 symposium on Eye tracking research and applications, ACM, New York, pp. 69–72.
- [43]. Partala, T., Surakka, V., 2003. Pupil size as an indication of affective processing, *IJHCS*, 59 (1–2), pp 185–198.
- [44]. Valverde, L, 2010, Inferencing emotions through the triangulation of pupil size data, facial heuristics&self-assessment techniques, proc: International conference on Mobile, Hybrid and Online learning, Vol. 62, pp. 972–988.
- [45]. Klingner, J., 2010. Fixation-aligned Pupillary Response Averaging. Proc: Symposium on Eye Tracking Research and Applications, ACM, New York, pp. 275–282.
- [46]. Just, M.A., Carpenter, P.A., 1993. The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal Experimental Psychology*, 47 (2), 310–339.
- [47]. Bacs, R.W., Walrath, L.C., 1992, Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics*, 23 (4), 243–254.
- [48]. Porter, G., Troscianko, T., Gilchrist, I.D., 2007. Effort during visual search and counting: Insights from pupillometry, *Quarterly Journal of Experimental Psychology*. 60 (2), 211–229.
- [49]. Kaklauskas, A., et al., 2012, Student progress assessment with the help of an intelligent pupil analysis system, *Engg Appl of AI*, Vol 26 no 1, pp 35-50.
- [50]. Chih-Ming Chen, Ling-Jiun Duh, 2008, Personalized web-based tutoring system based on fuzzy item response theory, *Expert Systems with Applications*, 34(4), pp 2298-2315.
- [51]. C. M. Chen, L. J. Duh, 2008, Personalized web-based tutoring system based on fuzzy item response theory, *Expert Systems with applications*, Vol 34, no. 4, pp 2298-2315.
- [52]. O. Ozyurt., 2013, Design and development of an innovative individualized adaptive and intelligent e-learning system for teaching–learning of probability unit: UZWEBMAT, *Expert Systems with Applications*, Vol 40 no 8, pp 2914-2940.
- [53]. Jay, T.B, 1983, The cognitive approach to computer courseware design and evaluation, *Educational Technology*, 23(1), pp 22-26.
- [54]. Jonassen, D.H. and Hannum, W.H, 1987, Research-based principles for designing computer software, *Educational Technology*, 27, pp 7-14.
- [55]. Larsen, R. E, 1985, What communication theories can teach the designer of computer-based training, *Educational Tech.*, 25(7), pp 16-19.
- [56]. O. Mich, E, 2013, Interactive stories and exercises with dynamic feedback for improving reading comprehension skills in deaf children, *Comp. & Education*, Vol 65, pp 34-44.
- [57]. Sylvia Encheva, On Providing Automated Support to Students Based on Their Personal Preferences, *IJCSI*, Vol. 10, Issue 2, No 3, March 2013.
- [58]. M. Goyal, E-learning: Current State of Art and Future Prospects, *IJCSI*, Vol. 9, Issue 3, No 2, May 2012.

Mitigate Black Hole Attack In Dynamic Source Routing (DSR) Protocol By Trapping

K.Mahamuni^{1*} and Dr.C.Chandrasekar²

^{1*} Assistant Professor, Department of Computer Applications, Government Arts College(Autonomous), Salem-7, TamilNadu, India.

² Associate Professor, Department of Computer Science, Periyar University, Salem-11, TamilNadu, India.

Abstract

Ad hoc network maximize the total network throughput by using all available nodes for routing and forwarding. MANETs are highly vulnerable to attacks than wired networks due to the open medium, dynamically changing network topology, cooperative algorithms, and lack of centralized monitoring. Hence, a node can misbehave and fail to establish route or route the data due to its malicious activity to decrease the performance of ad hoc network. In this paper, a method to mitigate the malicious nodes forming black hole attack in dynamic ad hoc network is proposed. This paper studies black hole attack impact in ad hoc networks with DSR routing protocol when the nodes are mobile. The proposed routing is based on DSR and is modified with detection algorithm. It is divided into two phases: Detection before route establishment and avoidance of malicious nodes during data forwarding. The silent feature of proposed scheme is its simplicity and effectiveness in detecting malicious nodes.

Keywords: Ad hoc network, network attacks, Dynamic Source Routing (DSR), Black Hole Attack, Trap Header(TH).

1. Introduction

Mobile Ad hoc Networks (MANET) are utilized to set up wireless communication in improvised environments without a predefined infrastructure or centralized administration [1]. MANET has been normally deployed in adverse and hostile environments where central authority point is not necessary. Another unique characteristic of MANET is the dynamic nature of its network topology which would be frequently changed due to the unpredictable mobility of nodes [2]. Furthermore, each mobile node in MANET plays a router role while transmitting data over the network. Hence, any compromised nodes under an adversary's control could cause significant damage to the functionality and security of its network since the impact would propagate in performing routing tasks.

Because these networks are temporary, they can be attacked from within, due to being constructed without protection, in poor conditions. Attacks are also launched if nodes are compromised. Another issue is the node number. Hundreds/thousands of nodes might be required in a network and security measures undertaken must be efficient and cost-effective for a vast network. Exchange of topological information among nodes is facilitated by routing protocols to establish routes and this is used by attackers for acts including bogus routing, incorrect forwarding, lack of error messages, restricted reply time, thereby leading to retransmission and inefficient routing[3]. Several work addressed the intrusion response actions in MANET by isolating uncooperative nodes based on the node reputation derived from their behaviors. Such a simple response against malicious nodes often neglects possible negative side effects involved with the response actions. In MANET scenario, improper countermeasures may cause the unexpected network partition, bringing additional damages to the network infrastructure. To address the above-mentioned critical issues, more flexible and adaptive response should be investigated.

Common attacks faced by networks include blackhole, grey hole and wormhole attacks, and IP spoofing[4]. Black hole attacks are malicious nodes that refuse to forward traffic[5]. External attacks can typically be prevented by using standard security mechanisms such as firewalls, encryption and so on. Internal attacks are typically more severe attacks, since malicious insider nodes already belong to the network as an authorized party and are thus protected with the security mechanisms the network and its services offer. Thus such malicious insiders who may even operate in a group may use the standard security means to actually protect their attacks. These kind of malicious parties are called compromised nodes, as their actions compromise the security of the whole ad hoc network.

In black hole attack, a malicious node uses its routing protocol in order to advertise itself for having the shortest path to the destination node or to the packet it wants to intercept. This hostile node advertises its availability of fresh routes irrespective of checking its routing table. In this way, attacker node will always have the availability in replying to the route request and thus intercept the data packet and retain it. In protocol based on flooding, the malicious node reply will be received by the requesting node before the reception of reply from actual node; hence a malicious and forged route is created. When this route is establish, now it's up to the node whether to drop all the packets or forward it to the unknown address [6, 7].

The method how malicious node fits in the data routes varies. Figure 1 shows how black hole problem arises, here node "A" want to send data packets to node "D" and initiate the route discovery process. So if node "C" is a malicious node then it will claim that it has active route to the specified destination as soon as it receives RREQ packets. It will then send the response to node "A" before any other node. In this way node "A" will think that this is the active route and thus active route discovery is complete. Node "A" will ignore all other replies and will start seeding data packets to node "C". In this way all the data packet will be lost consumed or lost.

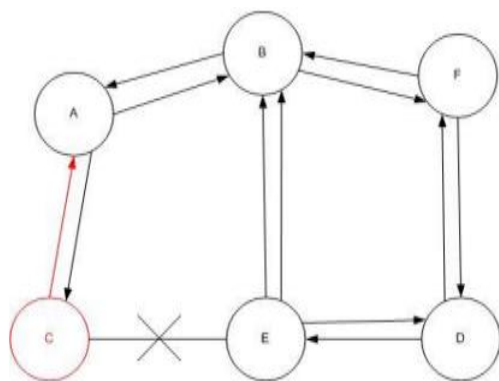


Figure 1: Black Hole Attack

In this paper, the proposed routing is based on DSR and is modified with detection algorithm. It is divided into two phases: Detection during route establishment and Detection during data forwarding. The silent feature of proposed scheme is its simplicity and effectiveness in detecting malicious nodes even when the network is highly dynamic. The rest of paper is organized as follows: Section 2 deals with the related works, Section 3 describes the methodology, Section 4 details the results and discussion, and section 5 concludes the paper.

2. Related Works

Hu et al.[8] presented a new protocol 'Ariadne' based on the DSR protocol for routing protection. Several authentication mechanisms such as digital signatures, MACs computed with pair-wise secret keys, or TESLA could be used with the proposed protocol. Hash chains are used to authenticate every route request protecting the network from overload, thus denial of service attacks are prevented. Attacks from compromised nodes from tampering with the uncompromised nodes are also prevented by the proposed method. Combinations of TESLA authenticators (MACs) are added by intermediate routers and a hashing technique to protect the discovered routes. The proposed method's security mechanisms are effective and can also be applied to wide variety of routing protocols.

Bhalaji et al.[9] analysed the black hole and cooperative black hole attack which is one of the new and possible attack in ad hoc networks. In this attack a malicious node advertises itself as having the shortest path to the node whose packets it wants to intercept. To reduce the probability it is proposed to wait and check the replies from all the neighbouring nodes to find a safe route. If these malicious nodes work together as a group then the damage will be very serious. This type of attack is called cooperative black hole attack. Our solution discovers the secure route between source and destination by identifying and isolating black hole nodes. In this paper, via simulation, the proposed solution are evaluated and compared it with standard DSR protocol in terms of throughput, Packet delivery ratio and latency.

Dadhania et al[10] investigated the performance of AODV and DSR in presence of black hole attack (malicious node) and without black hole attack with CBR (Constant Bit Rate) traffic under different scalable network mobility. Simulation was conducted to evaluate the effect and compare it with standard protocol in terms of throughput, Packet delivery ratio and End to End Delay. Extensive experiments using the network simulator-2 for 50 node ad hoc network was conducted. Results show that the AODV is more vulnerable to Black Hole attack than DSR.

In DPRAODV (Detection, Prevention and Reactive AODV) [11], they have designed a novel method to detect black hole attack: DPRAODV, which isolates that malicious node from the network. The agent stores the Destination sequence number of incoming route reply (RREPs) packets in the routing table and calculates the threshold value to

evaluate the dynamic training data in every time interval as in [12]. The solution makes the participating nodes realize that, one of their neighbors is malicious; the node thereafter is not allowed to participate in packet forwarding operation. In normal AODV, the node that receives the RREP packet first checks the value of sequence number in its routing table. The RREP packet is accepted if it has RREP_seq_no higher than the one in routing table. DPRAODV does an addition check to find whether the RREP_seq_no is higher than the threshold value. The threshold value is dynamically updated as in every time interval. As the value of RREP_seq_no is found to be higher than the threshold value, the node is suspected to be malicious and it adds the node to the black list. As the node detected an anomaly, it sends a new control packet, ALARM to its neighbors. The ALARM packet has the black list node as a parameter so that, the neighboring nodes know that RREP packet from the node is to be discarded. Further, if any node receives the RREP packet, it looks over the list, if the reply is from the blacklisted node; no processing is done for the same. It simply ignores the node and does not receive reply from that node again. So, in this way, the malicious node is isolated from the network by the ALARM packet.

Mittal et al[13] developed the routing protocol which can deal with single and cooperative black hole attack and also without degrading the performance of such networks. The approach provides a feasible solution for the Black hole node detection within the network by making use of "reputation tables" and assigning reputation values to the participating nodes. It improved the routing security of the existing association based DSR protocol using the concept of reputation of nodes value, by identifying and isolating black hole nodes working in a group.

3. Methodology

3.1. Dynamic Source Routing (DSR)

The Dynamic Source Routing (DSR) protocol is a on-demand routing protocol. DSR protocol maintains the route cache to store the route to the mobile node it is aware [14, 15]. This protocol composed of two major phases: route discovery and route maintenance. Whenever any node has the data to send, first it checks the route cache for the route to the destination. If it has the unexpired route, then it use it otherwise initiate a route discovery process by broadcasting the RREQ (Route Request) packet which contains the source

address and destination address. Whenever any intermediate node receives the RREQ, and it does not have the route to the destination it adds its own address in the route record and forward to its neighbor. RREP (Route Reply) is generated whenever RREQ reaches to destination node or intermediate node which has the route to destination in its route cache. Route maintenance mechanism is used to detect whether the path to the destination exist or not. Route maintenance uses the route error message and acknowledgement. Route error (RERR) message is initiated whenever the destination's data link layer recognize any transmission error.

DSR is suited for small to medium sized networks as its packet overhead can scale all the way down to zero when all nodes are relatively stationary [16]. The packet data overhead will increase significantly for networks with larger hop diameters as more routing information will need to be contained in the packet headers. The DSR protocol is composed of two main mechanisms that work together to allow the discovery and maintenance of source routes in the ad hoc network:

- **Route Discovery** is the mechanism by which a node S wishing to send a packet to a destination node D obtains a source route to D. Route Discovery is used only when S attempts to send a packet to D and does not already know a route to D.

- **Route Maintenance** is the mechanism by which node S is able to detect, while using a source route to D, if the network topology has changed such that it can no longer use its route to D because a link along the route no longer works. When Route Maintenance indicates a source route is broken, scan attempt to use any other route it happens to know to D, or it can invoke Route Discovery again to find a new route for subsequent packets to D. Route Maintenance for this route is used only when S is actually sending packets to D.

3.2. Proposed modification in the DSR

The proposed routing is based on DSR with modification for detection of black hole attack. It is divided into two phases: Detection before route establishment and avoidance of malicious nodes during data forwarding. The salient feature of proposed scheme is its simplicity and effectiveness in detecting malicious nodes in dynamic scenarios.

This algorithm has been designed based on the concept that malicious node may drop the packet or modify the packet. The DSR is modified to contain new header called Trap Header (TH). During detection phase, the nodes first sources the entire two hop neighbor node id's and sends trap packet with TH consisting of invalid data destination to its two hop neighbors. If the receiving node states that it has the route to the invalid destination in its cache, and has forwarded the data packet to next hop then the node is assumed to be a black hole malicious node. This information about the maliciousness is stored in the nodes. During route discovery, the nodes cross check the routes in its cache and if the route consists of a malicious node, the node invalidates that route and starts a fresh route discovery avoiding the malicious node. Thus, the proposed mechanism mitigates the black hole attack by a simple mechanism of trapping the malicious nodes and avoiding it in any of the routes during transmitting data packets.

4. Results and Discussion

The proposed DSR is simulated to evaluate its performance and compared with traditional DSR. The experiments are conducted for varying speed of the mobile nodes. The speed is varied from 10 Kmph to 90 Kmph and studied for the network performance. The black hole attack misbehaviour is defined as either drop the packets or not to forward the packet in the specified time interval. DSR routing protocol parameters were set as shown in Table 1.

Table 1: DSR Routing Parameters Used

Parameters	Values
Route expiry time	300 second
Request table size	64
Maximum transmission attempt	16
Timeout value for non-propagating requests	0.03 second
Gratuitous route reply timer	1 second
Maintenance hold off time during route maintenance	0.25 second
Maintenance acknowledgement time	0.5 second

Several performance metrics are used to compare the proposed DSR protocol with the existing one.

The following metrics were considered for the comparison were

Packet Delivery Ratio(PDR): It is the ratio of the number of packets received and the number of packets sent.

Average End to End delay: It gives the mean time (in seconds) taken by the packets to reach their respective destinations.

Table 2(a, b and c) tabulates the Number of hops to destination, end to end delay and packet delivery ratio obtained for the proposed DSR and DSR. Figure 2 to Figure 4 shows the same.

Table 2: Results of the experiments

Table 2(a): Values of No. of hops to destination

Mobility	No of hops to destination	
	DSR	Proposed DSR
10 Kmph	2.7	2.9
30 Kmph	3.2	3.6
50 Kmph	3.5	3.8
70 Kmph	3.9	4.1
90 Kmph	4.2	4.4

Table 2(b): Values of the end to end delay

Mobility	End to End Delay	
	DSR	Proposed DSR
10 Kmph	0.0514	0.0464
30 Kmph	0.0608	0.0582
50 Kmph	0.0684	0.0618
70 Kmph	0.0726	0.0638
90 Kmph	0.0784	0.0692

Table 2(c): Values of Packet Delivery Ratio

Mobility	Packet Delivery Ratio (PDR)	
	DSR	Proposed DSR
10 Kmph	0.9278	0.9432
30 Kmph	0.9148	0.9326
50 Kmph	0.8842	0.9014
70 Kmph	0.8621	0.8942
90 Kmph	0.8544	0.8824

It is observed that the number of hops for the proposed DSR is slightly more than the DSR as shown in figure 2. This is due to avoiding the malicious nodes in the network, while transmitting data packets to the destination. As the increase is negligible, when compared to DSR, the increase can be ignored.

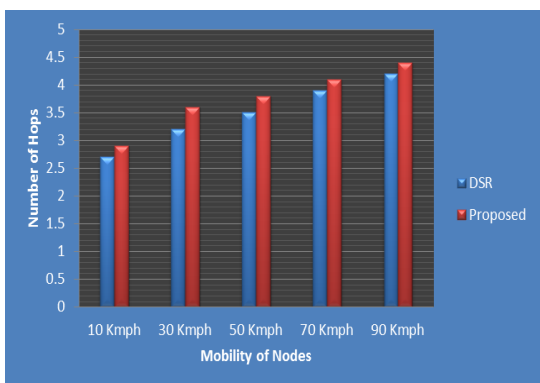


Figure 2. Number of Hops to Destination

The end to end delay in the proposed DSR is considerably less and it is observed that with the increase in number of nodes, the delay in DSR increases by 13.3%. Though, the number of hops from the source to destination increases, the end to end delay is less in the proposed DSR as shown in figure 3.

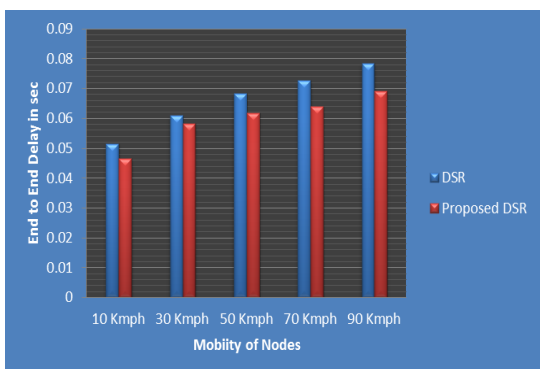


Figure 3. End to End Delay

The PDR improves with the use of modified DSR in the range of 1.57 % to 3.28% as shown in figure 4. It is observed from the tables and figures that the proposed DSR performance better than DSR in the presence of black hole attack.

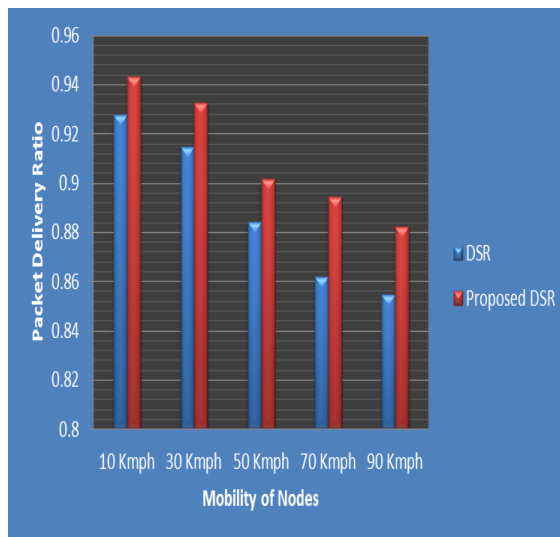


Figure 4. Packet Delivery Ratio

5. Conclusion

MANET networks are systems of mobile ad hoc networks which are presented dynamically and self-organized in temporary topologies. Internal attacks are typically more severe attacks, since malicious insider nodes already belong to the network as an authorized party and are thus protected with the security mechanisms the network and its services offer. The DSR routing is modified to include a Trap Header to identify malicious nodes. Experimental results demonstrate that the proposed DSR performance better than DSR in the presence of black hole attack under dynamic conditions.

References

- [1] Zhao, Z., Hu, H., Ahn, G. J., and Wu, R. Risk-aware response for mitigating MANET routing attacks. IEEE Conference In Global Telecommunications (GLOBECOM 2010), December 2010, pp. 1-6.
- [2] Mohapatra, P., Li, J., and Gui, C. QOS in Mobile Ad hoc Networks. IEEE Wireless Communications, June 2003, 10(3), pp. 44-52.
- [3] Buchegger, S., Boudec, J.-Y.L. Nodes bearing grudges: Towards routing security, fairness, and robustness in Mobile ad hoc networks, In Proceedings of the Tenth Euromicro Workshop on Parallel, Distributed and Network-based Processing, IEEE Computer Society, Canary Islands, Spain, January 2002, pp. 403-410.

- [4] Padmavathi, D. G., and Shanmugapriya, M. A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks. International Journal of Computer Science and Information Security(IJCSIS), 2009, 4(1&2), pp.1-9.
- [5] Papadimitratos, P. and Haas, Z.J. Securing the Internet Routing Infrastructure, IEEE Communications Magazine, October 2002, 40(10), pp. 60-68.
- [6] BrachaHod, Cooperative and Reliable Packet-Forwarding On top of AODV, www.cs.huji.ac.il/~dolev/pubs/reliable-aodv.pdf, 2005.
- [7] Sanjay Ramaswamy, Huirong Fu, ManoharSreekantaradhya, John Dixon and Kendall Nygard. Prevention of Cooperative Black Hole Attack in Wireless AdHoc Networks, In Proc. of International Conference on Wireless Networks, 2003.
- [8] Hu, Y. C., Perrig, A., and Johnson, D. B. Ariadne: A Secure On-Demand Routing Protocol for Ad hoc Networks. Wireless Networks, 2005, 11(1-2), pp. 21-38.
- [9] Bhalaji, N. and Shanmugam, A. Defense Strategy Using Trust Based Model to Mitigate Active Attacks in DSR Based MANET. Journal of Advances In Information Technology, May 2011, 2(2), pp. 92-98.
- [10] Dadhania, P., and Patel, S. Performance Evaluation of Routing Protocol like AODV and DSR under Black Hole Attacks. Performance Evaluation, International Journal of Engineering Research and Applications(IJERA), 2013, 3(1), pp. 1487-1491.
- [11] Payal N. Raj, Prashant B. Swadas. DPRAODV: A Dyanamic Learning System Against Blackhole Attack In AODV Based MANET. IJCSI International Journal of Computer Science Issues, 2009, Vol. 2, pp. 54-59.
- [12] Satoshi Kurosawa, Hidehisa Nakayama, Nei Kat, Abbas Jamalipour, and Yoshiaki Nemoto, "Detecting Blackhole Attack on AODV-based Mobile Ad Hoc Networks by Dynamic Learning Method", International Journal of Network Security, Nov. 2007, 5(3), pp. 338-346.
- [13] Mittal, S., and Taluja, H. Analysis of Cooperative Black Hole Attack Using Dynamic Source Protocol. International Journal of Advanced Research in Computer Science and Software Engineering(IJARCSSE), August 2012, 2(8), pp. 139-142.
- [14] Hu, Y., and Maltz, D. The Dynamic Source Routing Protocol (DSR) for Mobile Ad hoc Networks for IPv4. RFC 4728, February 2007, pp. 2-100.
- [15] Broch, J., Johnson, D. B., and Maltz, D. A. The Dynamic Source Routing protocol for Mobile Ad hoc Networks, 1998.
- [16] Bouhorma, M., Bentaout, H., and Boudhir, A. (2009, April). Performance comparison of ad-hoc routing protocols AODV and DSR. International Conference on Multimedia Computing and Systems'2009(ICMCS'09), 2-4 April 2009, pp. 511-514.

Bio data for Authors



Mr. K. Mahamuni received the B.Sc. (Computer Science) degree from the Bharathidasan University Tiruchirappalli, in 1992. He received his MCA and M.Phil (Computer Science) Degrees from Manonmaniam Sundaranar University, Tirunelveli in 1995 and 2001, respectively. He is working as an Assistant Professor,

Department of Computer Applications, Government Arts College (Autonomous), Salem, TamilNadu, Salem, India. He is currently pursuing the Ph.D. degree with the Department of Computer Science, Periyar University, Salem, TamilNadu, India. His research interests include Mobile Computing and Wireless Networking.



Dr. C. Chandrasekar received his Ph.D., degree from Periyar university. He is working as an Associate Professor, Department of Computer Science, Periyar University, Salem, TamilNadu, India. His areas of interest include Wireless networking, Mobile Computing, Image Processing,

Data mining, Computer Communications and Networks. He is a research guide at various universities in India. He has published more than 80 technical papers at various National & International conferences and 50 journals. He is a senior member of ISTE and CSI.

Transformation from E-commerce to M-commerce in Indian Context

Sujata P. Deshmukh¹, Prashant Deshmukh², G.T. Thampi³

¹ Mumbai University, India
Department of Computer engineering,
Lokmanya Tilak college of engineering, koprkhairane, Navi Mumbai, Maharashtra, India, 400709.

² Mumbai University, India
Department of mechanical engineering
Datta Meghe College of engineering, Airoli, Navi Mumbai, Maharashtra, India, 400708.

³ Mumbai University, India
Department of mechanical engineering
Thadomal Sahani College of engineering, Mumbai, Maharashtra, India, 400050.

Abstract

Electronic commerce and mobile commerce are not only frontier for doing global business and trade but also offers multiple benefits to the business, government and consumers on massive scale. Many companies, organizations, and communities in India are doing business using E-commerce and also are adopting M-commerce for doing business. But still E-commerce is not ubiquitous and pervasive and M-commerce is at nascent stage in India. The tremendous growth in mobile phone penetration shows that major Indian population has adopted to mobile phone and advancement in mobile technology and its usage is not limited for making basic phone calls, but can be used virtually in every sector of human activity—private, business, and government.

In this paper researchers attempted to answer the research question that 'Is India moving from E-commerce to M-commerce? For this study information is collected from TRAI, RBI guidelines available online and different websites, opinion given by top management of online companies, blogs and The Economic Times articles. In this study researchers identified the potential of E-commerce and M-commerce by reviewing its current status in India and considering the online users and their usage behavior. Through this paper, researchers attempt to identify the issues for the future growth of E-commerce and M-commerce with its present volume transactions in India and highlighted critical challenges before E-commerce and M-commerce would become an asset for common people. Researchers observed that customized and innovative services, right regulation and right models will drive the future E-commerce and M-commerce and it will occupy large segment in Indian market.

This paper evaluates the position of E-commerce and M-commerce in India which will be further helpful to for increasing productivity in India. At the same time it will also useful to design and implement different models of mobile commerce in India as model for the commerce in India.

Keywords—*Electronic Commerce, E-commerce, Electronic commerce in India, M-commerce in India, online users and usage behavior in India, Challenges of E-commerce in Indian market, Challenges of M-commerce in Indian market.*

1. Introduction

E-commerce, stands for electronic commerce, on the internet, it pertains to a website, which sells products or services directly from the site using a shopping cart or shopping basket system and allows payments through cards, e-banking, cash on delivery. Customers can purchase anything right from a Insurance Policy to pen o sitting comfortably in their office or home and gift it to someone sitting miles apart just by click of a mouse. It offers several benefits to businesses like easily reaching to a fast growing online community, providing unlimited shelf place for products and services, merging the global geographical and time zone boundaries and helping to reach national and global markets at low operating costs. Ease of internet access and navigation are the critical factors that will result in rapid adoption of E-commerce. Safe and secure payment modes are essential to popularize E-commerce in India. Though it offers many benefits to users, there are many reasons for not shopping online like are lack of trust, security concerns, uncertainty about product and service quality, delay or non-delivery of goods, and lack of touch-and-feel shopping experience [1][3][7][5][9].

Mobile Commerce (M-commerce) is the subset of electronic-commerce (e-commerce), which includes all e-commerce transactions, carried out using a mobile (hand held) device. M-commerce is the way of doing business in a state of motion [5][6][9]. M-commerce

depends on the availability of mobile connectivity. We have become used to making mobile phone calls anywhere, at any time and m-commerce is build on that capability. According to IT experts, in the future, consumers will be able to shop using a handheld computing device, PDA, wearable computer, mobile phone or smart devices. In virtually any place – malls, restaurants, hotels, airports and other locations – this user will be able to receive coupons, download information, receive sales offers, and perform credit card transactions. In such future scenarios, M-commerce means that customers can shop anywhere, anytime [29]. M-commerce offers multiple advantages like Ubiquity, personalization, flexibility, and distribution, instant connectivity, immediacy. There are many ways in which business, government and citizens of India could benefit from M-commerce like

1. Selling a product or service which is information based(delivery directly to mobile devices) or location based
2. Improving productivity by gathering time critical information(reports, photographs) and SMS based up-to-date information
3. The ability to access information on mobile, at affordable cost can change people's lives and livelihoods in rural areas. (Latest on the weather report or health services). It can be used as the medium to educate and create awareness among the rural people. Usages of Internet on mobile devices have lead to information access overcoming geographical barriers and removed the training cost of mobile technology.

2. Challenges of E-commerce and M-commerce in India

There has been tremendous growth in E-commerce and wireless technology in last decades. This has changed how people do the business with technology environment. For creating flexile and secure E-commerce and m-commerce in we need to leverage new technologies with old technologies and social and financial and technology development. The technology development issues that remained to be solved like extending handset battery life, increasing small screen usability, Less Powerful Processors and improving chip integration. M-commerce is a subset of E-commerce; it also faces some of the challenges of E-commerce like India has unbanked population, less credit card population, lack of fast postal services in rural India. The challenges of M-commerce in India are listed as below

1. Accessing the Internet by mobile phones is currently hindered down by slow transmission speeds, frequent disconnects, cost of Wireless connection

and wireless communication standards over which data is transmitted.

2. High-speed-bandwidth Internet connection not available to most citizens of the nation at an affordable rate.
3. Payment system that connects the utility of the mobile phone and the Internet together for the whole nation.
4. Limited Internet access among customers, lack of awareness about services and security among customers.
5. Lack of penetration of advance mobile device in rural area.[ss]
6. Multiple issues of trust in m-commerce technology, doubts about m-commerce security, and lack of widely accepted standards and lack of payment gateways (privacy of personal and business data connected over the Internet not assured; security and confidentiality of data not in place)
7. To deploy Ubiquitous IT Infrastructure and its maintenance: IT infrastructure can be defined as physical components such as hardware, software and network facility plus human components such as human expertise, manuals, and corporate culture constitute IT infrastructure of an organization. [4] [14] [15].

3. Status of E-commerce and M-commerce in India

Today E-commerce has become an integral part of our daily life. People use internet for various purposes which include: email, academic and financial information search, music and video on internet, chatting, online job search, gaming, booking railway/airline tickets, hotel reservation online news, internet telephony/video chat/voice chat, and online banking. E-commerce has touched the every field of human life from information search to entertainment, job search to matrimonial sites.

3.1 Online users and their usage behavior of E-commerce

India has witnessed an increase in the number of internet users. In India, the youth drive internet usage. School and college going students along with young men contribute to 72% of internet activity [11]. According to a report, jointly prepared by the Indian Council for Research on International Economic Relations (ICRIER), The Internet and Mobile Association of India (IAMAI) and the Department of IT India : Impact of Internet, a 10% increase in internet penetration can increase the GDP (gross domestic product) by 1.08 per cent. This will help the country add \$17 billion annually.

Research studies have indicated several factors responsible for the sudden growth of E-commerce in India such as:

1. It is predicted that by year 2015, around 65% of Indian population will be in the age group of 15-35 years. Since youth is an early adaptor of all technology, this seems to be a positive factor for the Indian E-commerce market.
2. India has an added advantage of lowest broadband prices by the organizations like BSNL. This enhances the potential of online transactions.
3. India has already started the efforts to provide biometric identity with unique identification number to all its citizens. Having this unique identity in place, and guidelines given by RBI for online transactions would make online financial transactions much safer as they can be easily tracked and subject to law [12].
4. Technology advancements such as VOIP have bridged the gap between buyers and sellers online. [33]
5. The emergence of blogs can be considered as an avenue for information dissemination and two-way communication for online retailers and E-commerce vendors.
6. Internet penetration is only about 10% (or about 121 million users) as against about 81% in the US and 36% in China. However this number continues to rise at a consistent pace because of falling prices for broadband connections. [32]
7. Innovation in e-commerce business models like no question asked return policies ranging from 7 days to 30 days, free product deliveries, the industry dynamics changing “cash on delivery” and product compare model offers benefits to customers than traditionally store [32]
8. Some companies have begun to develop support mechanisms for the entire cash on delivery model and are trying to reach at corners of India, including in the interiors where traditional logistics companies are still not completely present. The logistics companies are also shoring up their act and have started to build specific verticals and expertise to address the requirements of e-commerce companies[32]
9. Experts say that over the next 12-18 months there would be a couple of multi-product generalists who would be successful along with a leader in single product category.[32]
10. Secure online shopping model helped to increase e-commerce uptake. Currently only about 10 million people do online transactions out of an approximate population of 200 million credit and debit card holders. However the latest industry report by First Data Corporation and ICICI Merchant Services indicate that there are about 150 million users that are ‘ready’ for e-commerce. [32]
11. More importantly the report indicates that urban Indian consumers are now confident enough to make online purchases of up to US\$500 as against US\$40-100

in the recent past. So not only are the numbers of online shoppers projected to increase but there has been a real increase in the total value being spent online. [32]

12. According to CISCO report article in The Economics times, India will have highest traffic growth rate with 44% CAGR from 2012 to 2017 followed by Indonesia and South Africa. Monthly movement for fixed and mobile data is expected to reach about 121 Exabyte by 2017, from 44 Exabyte in 2012.
13. Improved fraud prevention and detection technologies have offered a safe and secure business environment and helped in preventing credit card frauds, identity thefts and phishing.
14. Consumers in cities are fast realizing the potential of the internet as a medium of transactions.
15. The young population finds online transactions much easier and safer [3].

3.2The status of mobile users and usage in India

A vast segment of the population that neither had a landline nor a bank account (unbanked) in their names but now not only they own a mobile handset but are also well balanced to transact on their mobile. The mobile channel has provided a rare opportunity not only to leapfrog years of poor infrastructure development but also in bypassing geographical constraints to bring massive benefits and lifestyle changes to millions of under-served people across India. India is a land of many languages but only 2 percent of the Indian population prefers reading in English. If the content is in local language, it will not only ensure quicker adoption by the user but also will be an instant success. Further, content developers are tempted to look at India as one market, and there are more players to share the pie. These apart, there exists the regional markets and a huge B2B market for m-commerce in India as well [4]. The status of mobile users and usage in India is as follows.

- There are currently about 867.80 million mobile subscribers with monthly growth 0.71% and this number is expected to touch 1.2 billion by 2015 according to the press release of TRAI, dated 29th May 2013, the country can easily expect M-commerce to take off faster than online or e-commerce and rising India as the world’s second largest mobile market after China.
- Total Broadband subscription reached 15.05 million in March 2013 from 15.00 million in February 2013. [31]
- Actively mobile Internet users are fairly young (18-35 years) and 81% percent users are using these devices as part of their daily routine - for sending emails, getting news and information, and shopping. More than half (61 percent) expect these devices to become universal payment tools.[4]

- Indians are also increasingly taking to mobile devices for not only search but shopping as well. The number of Smartphone users is rapidly increasing in India and with 4G services about to take off it's expected to get even more people going online. 27 million mobile users are active mobile internet users. More importantly, 20% users indicated intent to buy products through their mobile phones as against the current 4% and this number is expected to only increase in the next two to three years. [32]
- TechNavio's analysts forecast the Mobile Commerce market in India to grow at a CAGR of 71.06 percent over the period 2012-2016 and key factors contributing to this market growth is the growth in mobile subscriptions, government and regulatory support. However, the data security concerns of end-users could pose a challenge to the growth of this market.[34]
- This rise has resulted from introduction of voting based participative TV Programs, voting on some socio-economic-political issues in Newspapers; SMS based quiz/contests on mobile, MMS etc.[4] [11] [12].

This data shows that basic infrastructure required for Mobile communication has been built already. But there is a need for a major enhancement in ensuring the deployment of mobile Internet infrastructure for mobile commerce at affordable cost and desirable quality which can make infrastructure ubiquitous. Business can take the advantages of connected customers, can build brand digitally.

4.2 Emerging Services of M-commerce in India

- Increased mobile penetration and use of GPRS on handsets has resulted in the digital downloads market to cross Rs.2.55 Billion by end of March 2008.
- Monetary transfers or transactions through mobile phones are found to be much cheaper than traditional bank transfers as the transaction costs are much lower in the former.
- Innovative solutions for daily use such as payment of auto or taxi fares.
- Inter-bank transfers services are also on the anvil. ICICI Bank Ltd, India's largest private bank, has already started offering all its services through mobile phones (called mobile banking) since January 2008. Standard Chartered bank has launched a service that enables money transfer from any ATM to any mobile phone across the country. The recipient receives a pin number on his or her mobile phone and the sender has to convey the order number to the recipient. Using these two the

recipient can withdraw money from the bank's ATM's.

- Feasibility studies are being conducted to offer M-commerce to microfinance firms to enable them collect payments from remote areas [24].

4. Future of E-commerce and M-commerce in India

India is the seventh-largest by geographical area, the second most populous country in the world and third largest reservoir of technical human resource with fastest growing free market economy. Indian government has accepted free market policies in 1991, and globalization has increased trade and investments in Indian market. Indian E-commerce and M-commerce percentage is getting higher as more and more online retailers enter the market.[33]

4.1 E-Commerce Future Growth

The last few years have seen a rise in the number of companies' embracing E-commerce technologies and the Internet in India. Early E-commerce sites have been targeted towards the NRI's with Gift delivery services, books, Audio and videocassettes etc. but today E-commerce sites are available for every field of life with diverse products. People in smaller towns in India can access quality products and services similar to what people in the larger cities have access with the help of e-commerce sites. The growth of Indian e-commerce sector is US\$1.8 billion in 2012 and expected to reach 8.8 billion US \$ as per Forrester Research's projections for Asia-Pacific e-commerce sales by country and also notes e-commerce's rapid growth in India, where sales are forecast to grow 57% annually through 2016. Future of India's capital E-commerce market is as shown in table1

Table1: E-commerce market in India

Online Details	DEC 2012	DEC2016E	CAGR
Internet users	140 millions	390 millions	29%
Conversion to shoppers	9%	12%	NA
Online Shoppers	13 millions	47 millions	38%
Online shopping Market	\$809 million	\$3.62 billion	65%

E – Estimated

CAGR- Compound Annual Growth Rate

Source : The Economics Times dated Wednesday 10April2013.

As Indian E-commerce industry is growing, e-payments are increasing accessibility and proving convenience for customers as people are paying utilities, bills, tickets, travel, entertainment online instead of standing long queues. The driving forces for this are RBI policies for

secure online transactions with additional security code and 286 millions debit card penetration which is one fourth of population. The enhanced security and incentives, discounts, bonus points, cash back, preferential pricing provided by banks and merchants, can further enhance (add fuel) in E-commerce transactions.

Major Indian portal sites have also shifted towards E-commerce instead of depending on advertising revenue. The web communities built around these portal sites with content have been effectively targeted to sell everything from event and movie tickets, the grocery and computers. Most of the banks like ICICI bank, Global trust bank and UTI bank also have put their electronic banking over the internet facilities in place for the upcoming E-commerce market and the speed post and courier system has also improved tremendously in recent years with online package status at any moment in time. The future does look very bright for E-commerce in India with even the stock exchanges providing an online stock portfolio and status with a fifteen minute delay in prices [3][8][10].

Due to mobile and internet penetration and always on devices, rural customer is no longer rural customer like before, and lots of money is coming from government funding, income and transfer of funds, can open new market base for industries with wide range of challenges and frontiers. As knowledge and idea sharing increases, it shorten the product life cycle and rate of innovation is accelerated, and those can become asset for business.[ec] If people in India are made more aware, assurance of safe, convenient transactions, and help to save money will accelerate cash less economy and will benefit to merchants, consumers, overall economy of country.

B. M-Commerce Future Growth

Currently M-commerce in India is offering exciting and renewed services to the existing mobile users like pocket finder of Aircell, SMS based home appliances handling (ONIDA AC, FM), Projector within the mobile device. Social networking sites like Facebook, twitter are also available on mobile. But it has been limited primarily to basic banking transactions, purchase of travel tickets and payment of some utility bills, is finally taking off with banks, cellular operators and payment service providers coming together to find solutions that comply with regulatory guidelines. Security, a prime concern with banking regulators, has been an obstacle to the growth of M-commerce services such as the so-called mobile wallet, which helps make payments at retail outlets through text messages or Internet-enabled phones.

The Reserve Bank of India had issued guidelines against creation of the mobile wallet, a derivative of a phone firm's master bank account, from which millions of subscribers can be served on their mobile phones.

Californian mobile payment service provider Obopay Inc. and Bangalore's mChek India Payment Systems Pvt. Ltd are among a fast-expanding breed of mobile payment service providers which have planned offerings that work within the guidelines and bring the convenience of mobile transactions. Banks in India like ICICI Bank Ltd, SBI, Standard Chartered Bank have already started its Internet banking service on mobile phones. Bharti Telesoft Ltd provides software products and solutions to mobile operators, providing solutions for microfinance operations, especially in data collection work [26]. Data shows that currently products and services are merging as a result of cheap processing and widespread networking, and technologies are changing everything from our relationships to everyday objects. This has also changed the market trends and shopping decisions also. [33]

KPMG's study shows that consumers in India and China are leading the drive for personal banking and retail transactions via their mobiles. While unveiling the findings of the report, that 3G, 4G and BWA(Broadband Wireless Association) will further fuel the growth m-commerce in India as 3G, WiMax and the like that could support multimedia data delivery & these technologies should be able to ensure constant broadband speeds and seamless connectivity. Many service providers in India like BSNL, Airtel, Reliance, Tata Docomo, vodaphone are offering wireless 3G. The KPMG study has also revealed that in spite of the humongous growth of mobile telephony in India, more than half of mobile users did not wish to give up their landline. The reasons being, users felt landlines are more reliable than mobiles and it was the primary source for Internet access. This shows that in future E-commerce and M-commerce will occupy major share in Indian Market [25].

5. Conclusions

The potential of E-commerce is no more a matter of debate and India realizes the importance, as this is an emerging practice of businesses in today's world. The steps taken by RBI to improve E-commerce transactions and the data of online user and usage showed that India is adopting the E-commerce but it is not ubiquitous and pervasive due inability to leverage mobile and electronic commerce advances. Mobile penetration in India is very large but the M-commerce industry in India is in its nascent stage and is evolving every passing day. Data shows that many companies are coming forward in this field to offer new services. The huge unbanked population, the lack of credit-debit card penetration, higher charges on credit-debit usages and some barriers like lack of user trust and awareness in e-commerce and m-commerce technology, doubts about security and lack of widely accepted standards can little hinder the growth of m-commerce in India. But the worldwide acceptance and use of standards such as Japan's I-mode and

Europe's WAP, in combination with the work performed by market-based competition, collaboration of key-players, and regulations imposed by regulation authorities, securities in online payments are expected to boost consumer trust in M-commerce and strengthen its potential and perspectives. India has third largest reservoir of technical human resource, but it is not for medium of commerce for mass people, new models need to be developed and worked out with appropriate strategies to make electronic commerce and M-commerce as key policy for the development and progress in India. This current state will be further helpful to develop the new generation E-commerce i.e. mobile commerce for mass in India.

With the explosion of internet connectivity through mobile devices like Smartphone and tablets, millions of consumers are making decisions online and in this way enterprises can build the brand digitally and enhance productivity but government policies must ensure the cost effective methods/solutions. The advancements in technologies and innovative services shows that India is moving from E-commerce to M-commerce, and in future E-commerce and M-commerce will become asset for commerce by the people to the people in India.

REFERENCES

1. M. K. Harma , Dr. Ritvik Dubey, "Prospects of technological advancements in banking sector using Mobile Banking and position of India" International Association of Computer Science and Information Technology - Spring Conference, 2009, pp 291-295.
2. Preety Bareria, Mahesh Jain, Dr. D. K. Subramanian, Mahendra Pratap, "iMFAST - Journey so Far And Ahead", International Conference on Microelectronics, 2009,pp 141-144.
3. Pavan Soni, "M-Payment Between Banks Using SMS", IEEE, Vol. 98, No. 6, 2010, pp 903-905.
4. Deepali Sharma, Rajluxmi Murthy, D.Krishna Sundar, "Government Policies & Regulations: Impact on Mobile Commerce in Indian Context", 2nd International Conference on Mobile Governance, 2006.
5. White paper "A frame work for M_commerce", VeriSign, 2007, <http://www.verisign.com/in/static/037241.pdf>
6. Khawar Hameed, Hanifa Shah, Kamran Ahsan and Weijun Yang, "An Enterprise Architecture Framework for Mobile Commerce", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 3, No. 5, May 2010, pp 6-12, <http://www.ijcsi.org/papers/7-3-5-6-12.pdf>.
7. Yiming Xiang, Xiaobo Wu, Qi Chen , "Personal Innovativeness and Initial Adoption of M-Commerce: Toward an Integrated Model" IEEE ICMIT, 4th IEEE international conference, 2008, pp 652 - 657.
8. Amarjit Singh, M.P.Thapliyal, M.M.S.Rauthan, D.K.Joshi, "ENABLING E-COMMERCE IN INDIA "
9. TELE.NET magazine December, 2008
10. White paper, "Future of m-commerce", 2009, www.binarymantrasystems.com
11. Voice&Data Magazine, "The Essentials for The Success of M-Commerce in India", May 2010, <http://voicendata.ciol.com/content/ContributoryArticles/110051901.asp>
12. Business India 2.0 Forum, "M-Commerce India – The Future & Challenges faced!!", 29 August 2009, <http://ijsid.wordpress.com/2009/08/29/m-commerce-india-the-future-challenges-faced/>
13. Ook Lee, Dr. Woonghee Lee, "FACTORS THAT AFFECT SUCCESSFUL MOBILE COMMERCE", The International Association for Computer Information Systems (IACIS) ,2003
14. Ravi Tandon , Swarup Mandal and Debashis Saha, "M-Commerce-Issues and Challenges", 10th Annual International Conference on High Performance Computing (HiPC 2003)
15. Wolfgang König, "NATIONAL E-BUSINESS DIFFUSION", <http://www.wifrankfurt.de/publikationenNeu/NationalEbusine ssdiffusion.pdf>
16. BILL ANCKAR, DAVIDE D'INCAU, "VALUE CREATION IN MOBILE COMMERCE: FINDINGSFROM A CONSUMER SURVEY", JOURNAL OF INFORMATION TECHNOLOGY THEORY AND APPLICATION(JITTA), Vol. 4, Iss. 1, 2002
17. Pavan Soni, " M-Payment Between Banks Using SMS", IEEE,0018-9219,vol. 98, No. 6, June,2010
18. Shweta Sharma, Sugandha Mittal, "Prospects of E-Commerce in India", *ISCET 2010, ISBN 978-81-910304-0-2, pp no.43-48*
19. Sourabh Agarwal, Vikram Singh Mains, A.V. Chirputkar, Giri Hallur, "Mobile banking & m-Commerce and related issues", Symbiosis Knowledge Village,Pune
20. eCommerce in India: Overview and Reasons for Growth (http://www.india-reports.com/summary/ecommerce_in_india.aspx)
21. Online Users in India and Their User Behavior (http://www.india-reports.com/summary/ecommerce_in_india.aspx.)
22. ECommerce Industry in India – Government Regulations and Policies (<http://www.india-reports.com/summary/government-regulations-and-policies-towards-ecommerce-in-india-.mht>)
23. Source IAMA Report – 2008-2009(www.imbrint.com)
24. <http://www.indiareports.com/summary/Introduction to Mobile Commerce in India.mht>
25. M-commerce in India to gain more traction post 3G, BWA - CXO <http://www.cxotoday.com>
26. Mobile commerce gets ready for the next big step in India - Corporate News - http://www.livemint.com/livemint_com.mht
27. <http://www.roseindia.net/services/m-commerce-/mobile-commerce.shtml>
28. <http://www.rainbowskill.com/medium/fastest-3g-internet-in-india-provider-comparison.php>
29. <http://mobithinking.com/guide-mobile-web-India>
30. Report on online commerce, March,2011 <http://www.iamai.in>
31. Information Note to the Press (Press Release No. 38/2013) TRAI, New Delhi, 29th May, 2013, www.traigov.in
32. http://www.boston.com/business/blogs/global-business-hub/2012/03/the_sizzling_in.html
33. The Economic times articles.
34. <http://www.reportlinker.com/p01095417/Mobile-Commerce-Market-in-India-2012-2016.html>

Sujata P. deshमुख (B.E.(comp. Sci & engg.), M.E.(IT), Ph.D.(pursuing in IT) from Mumbai University, India. She is currently working as Assistant professor in LTCE, and her areas of interest are E-commerce and M-commerce, Ubiquitous computing.

Prashnant Deshmukh (B.E.(Production engg.), M.E.(Mech-CAD/CAM), Ph.D.(pursuing in Mech. Engg.) from Mumbai University, India. He is currently working as Assistant professor in DMCE, and his areas of interest are ERP, Production Engg, Material technology and CAD/CAM.

Dr. G. T. Thampi M.E.(Mech.), Ph.D.(Technology.) from Mumbai University, India. He is currently working as Principal of TSCE, and RRC member of Mumbai university, X-chairman od BoS(IT). His areas of interest are BPR, E-commerce and M-commerce, ERP.

An Expressive Role-based Approach for Improving Distributed Collaboration Transparency

Abdullah O. Al-Zaghameem¹ and Mohammad Alfraheed²

¹ Department of Mathematics and Computer Science, Tafila Technical University
Tafila, 66110, Jordan

² Department of Mathematics and Computer Science, Tafila Technical University
Tafila, 66110, Jordan

Abstract

This paper presents an approach for applying the Remote Role-Playing (RRP) concepts in the development of Distributed Collaborative Applications (DCAs) that use the collaboration-based (role-based) design. The paper studies a programming model called Object Teams (OT), which aims at implementing this design for the object-orientated languages. Then, it introduces the RRP for improving the DCAs modularity through mapping the fundamentals of the OT model to distributed environments. The approach is demonstrated through developing a simple case study.

Keywords: *Distributed collaborative application (DCA); DCA Modularity; Role-based design; Object Teams (OT)*

1. Introduction

A Distributed Collaborative Application (DCA) could be defined as a group of separated programs which are executed on several network nodes to achieve a shared goal. The demand to smoothly design and implement DCAs is highly increased as the complexity of the development of these applications has increased too. The main complexity in the development of DCAs is the decomposition of system functionalities into separated components [3]. Recently, the Aspect-Oriented Programming (AOP) proves itself as a promising technique to improve the quality of software by decreasing the level of code scattering and tangling [2] [4]. As a consequence, AOP concepts are applied in various software engineering fields. For instance, several approaches have been developed like Tako-approach [6] (a black-box approach) and AO-CVE - approach [1], to improve the modularity of the DCAs by separating the *collaborative functionalities (CFs)* of components from the application core functionality.

In the context of collaboration modularity, the Object-oriented Collaboration-based Design (OOCBD) describes a methodology for decomposing object-oriented applications into a set of classes and a set of collaboration modules [7]. The AOP-based approaches lack the appropriate module for expressing collaboration modules. One of the recent approaches that adopt the OOCBD is

Object Teams (OT) [5]. The OT programming model offers a Role-based technique [9] [10]. Role is one of the foundation concepts for building collaborative applications [8]. .

The OT model captures collaborations in new modules called *Teams*, and separates the collaborative functionality of application objects inside modules called *Roles* [5 and 14]. An object in a collaborative application participates by playing a role in the collaboration via Role-playing process [11]. The OT model has many features; the capability to decompose collaborative-based applications in modules in expressive and modular way. Moreover, offering an integration foundation between the AOP concepts and the Object-orientation. The main contribution of this research is to map the OT's features to distributed environments. The realization of this mapping improves the DCAs modularity and offers expressive representation of the CFs; hence simplifies DCA composition. In addition, the mapping assists developers to construct easy to develop, evolve, and maintain DCAs.

In this paper, the Remote Role-Playing (RRP) is introduced as an approach to enable the distributed object-oriented application components to play the roles of "teams" remotely (i.e. in distributed environment), while preserving the semantics of the OT role-playing. The team module in RRP (called *remote team*) is used to represent the context of the collaboration, which enables the dynamic management of collaboration activities like *activating/deactivating* objects participation. The role modules (called *remote roles*), on the other hand, are used to capture the CFs of application components in collaborations. The utilization of modularity of collaboration-based design in RRP facilitates resolving primary collaborative problems like conflict resolution [12] and early conflict resolution [13]; because remote roles *intercept* the CFs of application objects before they are employed in collaboration.

The paper is organized as follows: in Section 2, a simple collaborative application will be developed as a case study. Section 3 presents an overview of the OT model, and highlights its key features. Section 4 presents the RRP approach; which maps OT features to distributed environments. The case study of Section 2 will be used to

demonstrate the RRP. The section addresses the requirements of accomplishment the mapping process and emphasizes the features offered by RRP. In Section 5, an evaluation discussion will be held for the performance of RRP. Section 6 discusses the related works, and in Section 7 the conclusions.

2. Case Study: Simple Painting Collaboration

In this paper, the following simple collaborative scenario will be used to demonstrate the distributed collaboration based on remote role-playing: Consider a collaboration of multiple users for drawing a shared “painting”. Using the object-orientation model, the following primary application entities could be recognized: the painter (who performs drawing), the shapes which a painter can draw, and the place on which shapes could be drawn. In the world of patterns, the Model-View-Controller pattern fits best the design of this application. The painter (the Controller) works on a set of shapes (the Model). The controller can create and add shapes on a graphical interface (View). Fig. 1 illustrates the UML class diagram of this composition. Java programming language is used to implement the case study as shown in Fig. 2. The figure illustrates a code snippet of the implementation of **Painter** class.

Inspecting the code in Fig. 2, the method **paint** is recognized as a *Collaborative Functionality (CF)*. That is, it points out to the part of painter object’s behavior in the collaboration processes. Depending on object-oriented Collaboration-based Design (CBD), the “painting collaboration” *crosscuts* painters’ functionality at paint method. Fig. 3 shows the CBD, where **CollPainting** (the name of our painting collaboration) crosscuts **Painter** class.

In OT model, the intersection between **Painter** class and **CollPainting** collaboration is the *role* which **Painter** objects can play in that collaboration. The OT model captures the **CollPainting** collaboration in a *team* module. For modeling purposes, the OT model uses UML notions to model collaborations as an integration of package and class diagrams. Fig. 4 illustrates the OT representation of “collaborative painting” application. In

this way, the CF of **Painter** class is captured in the **CoPainter** role.

A Team module in OT/J [14] (the programming language implementing the OT model in Java) is a first-class entity. Thus, it can declare attributes and methods (see Fig. 4), extends other teams and, most importantly, it can be instantiated. Likewise, roles can have their own attributes and methods.

The OT model organizes the collaborative relationship between teams and application classes (called *bases*) through a high expressive relationship called *playedBy*. The “playedBy” relationship selects one base class to play exactly one role in any specific team. At runtime, it binds base objects to role instances by the mediation of team instance.

2.1 The Collaborating Process

To facilitate application base objects participation in teams, the “playedBy” relationship establishes two types of communication between bases and their roles: the first one is called *Callin Method Binding (CIMB)*, which enables base objects to *call into* role instances specific methods *after*, *before*, or *in replacement* of their methods. For example, the expression `{collPaint ← after paint;}`, shown in Fig. 4, is a CIMB that instructs **Painter** objects to call the method **collPaint** *after* they call their **paint** method. The method **collPaint** is named as role’s *callin* method. In fact, the OT model introduces the CIMB expression and role’s callins as counterparts to, respectively, *pointcut* and *advice* concepts of AOP languages like AspectJ [15].

The second communication type is called *Callout Method Binding (COMB)*, and indicates that a role instance declare a method, which is not available locally, by *calling out* to a method of the associated base object.

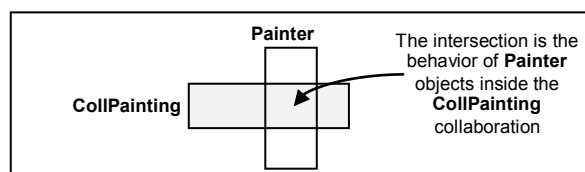


Fig.3 CBD of the collaborative Painting application

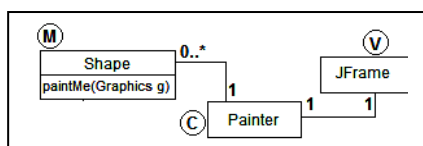


Fig.1. Class diagram of painting application

```

1. class Painter
2. implements MouseListener, ..
3. {
4.     JFrame window = ..;
5.     List<Shape> shapesList = ..;
6.     :
7.     public void prepareShape(Shape s){..}
8.     public void paint(Shape s){..}
9. }
    
```

Fig.2 Part of **Painter** class implementation in Java.

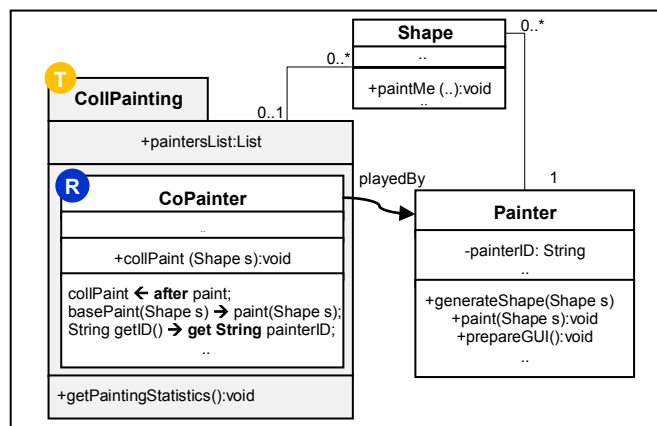


Fig.4 Modularizing the CBD of “painting collaboration” in OT.

OT presents two types of COMBs: the first is called method-COMB, which indicates that the stated role method will invoke a base method. For example, the expression {basePaint(Shape s) → paint(Shape s);} will dispatch the calls made to **basePaint** method (which is never implemented) toward the base method **paint**. The second COMB type is called field-COMB, which enables role instances to *get* or *set* the values of their base objects' fields. For example, the expression {String getID() → **get String painterID;**} will get the value of **painterID** field of the bound base object whenever a call is made to the method **getID**.

The CIMB and COMB mechanisms together form a complete communication channel between base objects and their roles, which facilitates control and data flow, and fulfills the needs of role and base objects for accomplishing precise collaborating.

2.2 Collaboration Programming

In OT/J, developers can define teams by using the keyword “team”, and define role classes as inner-classes. Fig. 5 shows a code snippet of the implementation of **CollPainting** team. Note the expressive binding between **Painter** base class and **CoPainter** role class through the keyword “playedBy” at line 3.

3. Obstacles to Remote Role-Playing in OT/J

The Remote Role-Playing (RRP) is the technique aims at enabling the objects of distributed applications to play the roles of OT/J applications *remotely* while preserving the semantics of the OT role-playing. In practice, the OT/J programming language involves the following specifications that technically prevent the remote role-playing to be applied:

- Obs-1: The OT/J's *weaver* adopts the mechanism of *Load-Time Transformation (LTT)* [14] to weave roles in the bytecode of application base classes. The LTT technique is widely used in AOP for injecting aspects into the application's

```

1. public team class CollPainting{
2.     List paintersList =...;
3.     protected class CoPainter playedBy Painter {
4.         ..
5.         public void collPaint(Shape s)
6.         {
7.             System.out.println("Painter:"+getID());
8.
9.             /* paint the shape s at all
10.              other participants painters */
11.             for(CoPainter coP : paintersList)
12.             if(coP != this)
13.                 coP.basePaint(s);
14.         }
15.     }
16.     collPaint <- after paint; // CIMB
17.     basePaint(Shape s) -> paint(Shape s); // COMB
18.     String getID() -> get String painterID; // COMB
19. }
20. public void getPaintingStatistics() {...}
21. }
    
```

Fig. 5 Implementation of **CollPainting** team in OT/J.

business logic. In general, the using of LTT results in hard coding roles' callin methods inside base classes' bytecode. This will prevent mapping OT/J applications to distributed environments because teams are taught to bind their roles to local base objects only. Similarly, role classes are compiled and transformed in a way the generated role instances deal with local base references only, which are irreplaceable by remote references at the source code.

- Obs-2: OT/J supports only *static* role-playing, i.e. a specific base object can play only those roles that woven into its base class. Thus, it cannot play new roles *dynamically* at runtime. Practically, this impacts the capability of OT/J in developing dynamic collaborations.
- Obs-3: In Java-based distributed applications, distributed components often represented via *contract-based* designs like CORBA-IDL [16] or *remote interfaces* of Java-RMI [17]. In both cases, OT/J does not fully support playing roles by bases that are *interface*; due to implementation limitations [14] (§2.1.1).

4. Mapping OT/J Applications to Distributed Environments

In this section, the conception of mapping OT/J applications to distributed environments is presented in order to investigate the shortcomings of OT/J to support the remote playing of roles. The section first addresses the requirements for a precise mapping (see section 4.2), and introduces the Remote Role-Playing (RRP) (see section 4.3).

4.1 Distributing OT/J Applications

To make the application of case study in Section 2 a true DCA, painters need to be able to participate in “painting collaboration” over a real distributed environment. The participation imposes the existence of **Painter** objects and **CollPainting** team (and its role instances) at separate network nodes or application processes. To demonstrate the impact of technical problems described previously (in section 2), three Painter-application instances are deployed on three different nodes (H1, H2, and H3) as shown in Fig. 6. In

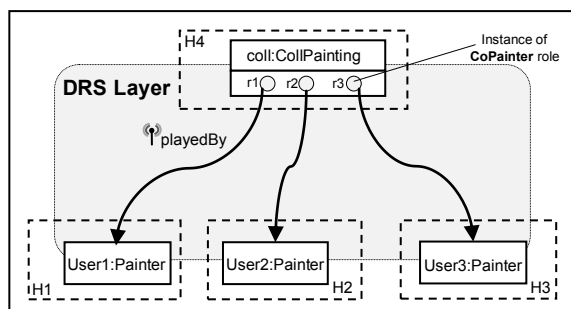


Fig.6 Deploying “Painting collaboration” on distributed Environment.

addition, an individual **CollPainting** team instance must be deployed at each node to capacitate the deployed **Painter** objects playing the **CoPainter** role. In this case, three-separated *local collaborations (teams)* must be created. The problem arises here is that “User1” (H1) is not be able to share the CF (i.e. painting) with “User2” (H2) and “User3” (H3) because she plays a local role.

Another problem shows up is that the state of separated local-collaboration instances (i.e. teams) must be synchronized all the time. For example, the activation of one of them should lead to activating others. The same thing must be done in case of collaboration deactivation; otherwise, collaboration inconsistency is encountered.

To overcome these problems, a single team instance has to be deployed and all **Painter** objects should play its roles remotely; thus, synchronizing team-instance copies is eliminated instead of preserving a unified consistent state among them. Furthermore, the inter-relationships between team’s roles are preserved (if any) unbroken. For these reasons, the RRP is introduced as a communication layer over which the deployed **Painter** objects can play the **CoPainter** role *remotely* (see Fig. 4). The “playedBy” relationship is marked with the *antenna symbol* to indicate that it binds discrete base and role objects. Having a single team instance is very important to unify the effects of role-playing on base objects via team *activation/deactivation* operations.

4.2 Requirements of RRP

The separation of base objects and team instance of **CollPainting** results in the broken of “playedBy” relationships. The breaking results due to hard coding roles inside base class’s bytecode by the OT/J transformers. The fundamental idea in the RRP approach is to *reformulate* base and team classes (including roles), which are involved in *remote* “playedBy” relationships. In such way, their objects are able to communicate over distributed environments. This communication should guarantee precise CIMB and COMB executions, and preserves the semantics of the local “playedBy” relationship. Therefore, the accomplishment of the following primary Requirements (R) is essential in order to realize that communication:

- **R1:** The participation of distinct base objects of class **Painter** in the collaboration. In this requirement, the capability of the base objects of class **Painter** deployed on hosts H1, H2 and H3 (hereafter called *remote base objects*) has to participate in the collaboration. In OT/J, this requirement implies that remote base object’s functionality must be remained intercepted by the role’s CIMBs. Moreover, remote base objects must determine the location of the team instance, which encloses their roles before any of their functionalities (i.e. methods) is invoked. Simultaneously, a remote base object must *dispatch* its CF remotely to the desired team instance; thus, a distributed collaboration could be achieved.

- **R2:** *creating* role instances based on demand. In this requirement, the capability of the **CollPainting** team instance deployed on host H4 (hereafter called *remote team instance*) has to *create* role instances on-demand. Therefore, the requirement involves its capability to bind the created role instances to the remote base objects.
- **R3:** addressing the declared COMBs on the exact remote base objects. In this requirement, the capability of role instances of **CoPainter** has to issue the declared COMBs on the exact remote base objects associated with. For example, the role instance “r1” (shown in Fig. 6) should invoke the field-COMB **getID()** (see lines 6 and 13 of Fig. 5) on the remote base object “User1” only.

4.3 The RRP Implementation

In terms of implementation of the RRP, two issues have to be taken in account. First, carrying out these requirements imposes the relaxation of tight coupling between base and team classes (including its *remote* roles) before they are loaded into the Java Virtual Machines (JVMs) for execution. Furthermore, requirements of RRP have to be fulfilled without violating the core functionality of base or team objects. In other words, remote base objects of **Painter** class have not to detect that they are playing remote roles, and they have to exhibit their behaviors, even though, for example, the team found deactivated. Second, the remote base objects must be affected by the role-playing to achieve the participation in the collaboration. To achieve these two issues, the (Distributed OT/J (DOT/J)) Runtime System (DRS) layer is here developed to first operate underneath the remote objects (bases and teams) and to second facilitate the RRP activities (see Fig. 6). Therefore, a need to communication layer is aroused. Two types of communications are introduced:

1) *An inter-communication between the DRS components:* The DRS layer comprises two main units (as shown in Fig. 7) which are the Group Communication Unit (GCU) and the Distributed Objects and Teams Manager (DOTM). The latter provides the necessary coordination for remote base objects and team instances in a distributed OT/J application to accomplish a precise binding. The Group Communication Unit (GCU), on the other hand, is responsible of the *transparent* deployment of team instances and the contextual information of roles on every application node. Moreover, GCU implements a reliable

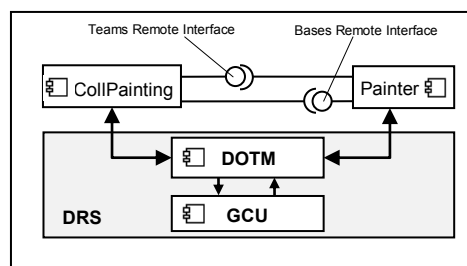


Fig.7 The RRP Runtime Infrastructure.

communication protocol between the deployed DOTMs, so that they can preserve the collaboration consistency by synchronizing their states; in particular, the *Remote Teams List (RTL)*.

2) *The actual RRP communication*: This type of communication involves the direct communication between remote base objects and their roles (via the remote team instance) to execute the CIMBs and COMBs declared in these roles. Establishing smoothly this type of communication requires that remote base objects and remote teams have to communicate remotely. In addition, the relaxation of base and role bytecodes needs to be accomplished. Therefore, the *Provided/Required Interface* design is adopted to glue the remote base objects and remote teams of application. The provided/required interface design nicely captures the relationship between remote objects. Thus, remote base objects of **Painter** must provide a specific remote interface, and require another interface from the remote team **CollPainting**. Likewise, the remote team instance of **CollPainting** provides a remote interface and requires the interface of **Painter** objects. Fig. 7 depicts this relationship.

In order to preserve the integrity of remote base and remote team objects, the structure of their classes after *gluing* must not be negatively violated, i.e. their hierarchical structures must remain untouched. Consequently, the LTT is adopted to *reformulate* remote base **Painter** and remote team **CollPainting** classes into *distributed components* as shown in Fig. 7. The transformation details are beyond the scope of this paper. Nevertheless, it is important here to mention that our bytecode transformers (see Obs 1 in section 3) *reuse* the transformed classes of the OT/J application, and carefully *replace* the features of local “playedBy” relationship with remote ones. For example, role instance “r1” (shown in Fig. 6) is bound to the *remote stub* generated from the remote interface provided by the remote base object “User1”.

4.4 The Engagement between Remote Base Objects and Remote Team Instances

Enabling remote base objects of **Painter** class to allocate dynamically the remote team instance “coll” requires that the team instance *registers* itself in the DOTM (i.e. DOTM runs on host H4). The registration process involves providing the local DOTM component with contextual information about the team instance such as its *remote stub* and a list of the CIMB expressions declared by its *remote* roles. As a consequence, the DOTM *broadcasts* the record of the registered team (via the GCU) to all the DOTMs executed on other application nodes. Moreover, the DOTMs deployed on hosts H1, H2 and H3 keep an identical RTL.

Once a registration record arrives to the local DOTM at host H1, it *notifies* all the remote base objects coexisting at the same node (in this case the instance

“User1”). At this moment, “User1” obtains the knowledge that a team instance comprises one of its roles. In OT/J, a base object starts playing a role *after* it has been bounded to an instance of that role. The binding process has been carried out when one of base object’s methods has been intercepted by a declared CIMB. Thus, as “User1” already obtains the recent RTL, it can check whether any of its functionalities has matched any of the CIMBs. To facilitate this task, the bytecode transformer injects a *trap* at every declared method in **Painter** class, and dispatches their calls to a central *dispatcher* method. The dispatcher method traverses the RTL and extracts those CIMBs matching the current invocation.

Here, two different transformation strategies can be chosen:

The first is to invasively trap all base methods (called *total hook weaving* [18]); thus, the dynamic playing of roles is supported at runtime whenever new teams are attached to the application without interrupting its execution. Within this strategy, the remote team instances and remote base objects are required to provide *generic remote interfaces*. As a consequence, remote base objects can *plug* into any dynamic collaboration (team), and different remote team instances can bind the same remote base object using a single remote stub. In addition, a transparent conversion of base objects into collaborative-objects is achieved.

The second strategy is to declare those base class functionalities desired for being intercepted by CIMBs in a data file (like XML). The latter is then directed to the transformer in order to target only these functionalities. In this way, a great performance is saved overhead at runtime. However, the desired dynamic evolution of the DCAs is restricted.

When “User1” draws a specific shape, the method **paint** is called and trapped by the dispatcher. The dispatcher, in turn, detects a CIMB matching. Consequently, the dispatcher extracts the accurate *Team Level Wrapper Method (TLWM)* of the remote team instance “coll” from the contextual information. In addition, the extraction guarantees that the TLWM corresponds to the role’s call in method **collPaint**. The latter then invokes that TLWM via the remote stub of “coll”. The team instance “coll” receives, as an argument to TLWM, the remote stub of “User1” base object. Therefore, it can create a new role instance (in our case it was “r1”) and binds it to the remote stub. According to the latter, the role instance “r1” can address COMBs on the remote base object “User1” using the RMI invocations.

5. Evaluation

To evaluate RRP, the application shown in Fig. 6 is executed as follows: first, the Team-application is run at host H4. Therefore the three Painter-applications are executed at H1, H2 and H3. The role **CoPainter** is implemented, so that it adds the participant painter to the painters list (see line 2 of Fig. 5). We achieve this by intercepting the method **prepareGUI** of **Painter** objects

by a CIMB (not shown in Fig. 5). After then, each one of the painters “User1”, “User2”, and “User3” paints randomly 10 shapes. The **CollPainting** team has been provided with a GUI, so it can clone the *shared painting* as well. The role’s callin method collPaint will take care of synchronizing the painted shapes of a specific painter at others GUI (lines 7 to 9 of Fig. 5). As expected, the 30 shapes appear on all Painter- and Team-applications’ GUIs.

For performance analysis, the runtime required for carrying out a complete CIMB interceptions is recorded, which includes (1) dispatching the control flow from remote base object to the bounded role instance, (2) broadcasting the painted shape to all other painters, and (3) painting that shape at the team’s GUI. The recorded values (in milliseconds) are shown in Table 1. The average runtime values give clear symptoms for a promising approach. However, further evaluation on more advanced case studies is one of our future works. The time required at the first interception in all cases is higher than the rest. The extra time is consumed by the Java-RMI system for generating dynamic proxy objects.

Table 1: Runtime values for intercepting **paint** method 10 times at the three painter-applications.

No.	User1 (H1)	User2 (H2)	User3 (H3)	Avg.
1	50	20	30	25
2	16	14	7	10
3	10	13	7	8
4	12	6	6	7
5	14	13	6	10
6	17	16	10	12
7	8	15	7	9
8	18	13	10	12
9	11	9	11	10
10	17	13	6	12
Avg.	17.3	13.2	10	11.5

6. Related Works

The RRP could be considered as a distributed-AOP approach; because it supports intercepting the remote base objects’ functionalities by the remote CIMBs of roles, and then executing the associated advices *remotely*. In addition, it could be introduced as an approach for improving the composition and modularity of DCAs. In this regard, the RRP shares several features with TaKo [6]. First, RRP and TaKo are AOP-approaches which address the transparent collaboration between legacy applications that were not designed for collaborating. TaKo proposes a full blackbox approach for supporting collaboration transparency. However, it is environment-specific as it targets AWT- and Swing-based Java applications only. In contrast, RRP presents more expressive approach to design and model the collaboration and the collaborating functionalities. After that, the RRP infrastructure operates to provide the necessary facilitation for accurate distributed collaboration. In RRP, the strategies of solving problems like Collision Detection, Priority, Logging, etc. are easy to customize and implement.

In [1], a proposal for employing the AOP concepts in the development of collaborative virtual environments (CVEs) is presented. The approach stands on intercepting the functionalities of application components, and dispatch control flow to a dedicated middleware layer, which interconnects between components and aspects dynamically. All aspects must be previously registered in this layer to ensure accurate interconnections. The RRP offers a similar approach; remote team instances need to register in the DOTM. However, the AO-CVE as an AOP-based approach lacks support for clear collaboration modularity and expressive relationships between collaborative components and aspects as in RRP.

Service-Oriented Architecture (SOA) is used in [19] besides a set of off-the-shelf technologies to develop a collaborative authoring application, which involves five users collaborating over the internet. This and other approaches like CoDesign [13] and GroupUML [20] are classified as collaborative software design and modeling environments. These and alike approaches cope mainly with problems of conflict detection, shared state and time synchronization, among others. Our approach offers better separation between the core functionality of application components and the collaborative functionalities each component exhibits in the collaboration. Even though SOA is (in concept) similar to team module of OT, teams provide twofold facilities; a clear contextual collaborations and providing services. In addition, the client-server topology depicted in base-team relationships capacitates team instances for resolving any CFs conflicts.

7. Conclusions and Future work

This paper has presented the Remote Role-Playing (RRP) concept as a promising approach for improving the modularity and composition of distributed collaborative applications (DCAs). The realization of this approach has been discussed through the process of mapping the fundamentals of the Object Teams (OT) model. The OT model implements the collaboration-based (role-based) designs for the object-oriented languages. A primary key feature of the OT model is that it presents a high expressive modularization technique through introducing collaborations as first-class entities; which makes RRP an approach that is neither total black-box nor total white-box.

The paper has presented a simple case study to demonstrate the approach and emphasize its capability for supporting transparent collaboration between legacy Java-based applications without their source code needed. In addition, it can offer several programming capabilities for DCAs like managing collaboration contexts via team instances, besides the dynamic participation in collaborations. That is, applications’ base objects can enroll in several collaborations dynamically. As future work, further evaluations need to be established in real case studies.

References

- [1] M. Pinto, M. Amor, L. Fuentes, and J. M. Troya, "Collaborative virtual environment development: An Aspect-oriented approach," In Proceedings of DDMA Workshop. Phoenix, Arizona, pp. 97-102, April 2001.
- [2] G. Kiczales, and *et al.*, "Aspect-oriented Programming," In European Conference on Object-Oriented Programming - ECOOP'97, 1997.
- [3] M. Pinto, L. Fuentes, J.M. Troya, "Towards an aspect-oriented framework in the design of collaborative virtual environments," In Proceedings of the Eighth IEEE Workshop on Future Trends of Distributed Computing Systems (FTDCS.01), 2001.
- [4] S. Subotic, and J. Bishop, "Emergent Behaviour of Aspects in High Performance and Distributed Computing," Proceedings of SAICSIT'05, pp. 11-19, 2005.
- [5] S. Herrmann, "Object Teams: Improving Modularity for Crosscutting Collaborations," In NetObjectDays Conference on Objects, Components, Architectures, Services, and Applications for a Networked World (NODE '02), Springer-Verlag, London, UK, pp. 248-264, 2002.
- [6] R. Mondéjar, P. García-López, E. Fernández-Casado, and C. Pairot, "TaKo: Providing transparent collaboration on single-user applications," Computer Languages, Systems and Structures, Elsevier Science Publishers, Amsterdam, The Netherlands, Vol. 38, Issue. 1, pp. 108-121, April 2012.
- [7] M. Mezini and K Lieberherr, "Adaptive plug-and-play components for evolutionary software development," In Proceedings of the 13th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications (OOPSLA '98). ACM, New York, NY, USA, pp. 97-116, October 1998.
- [8] H. Zhu, "Some issues of role-based collaboration," Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on , vol.2, pp. 687- 690, May 2003.
- [9] H. Zhu, "Role mechanisms in collaborative systems," International Journal of Production Research, Vol. 44, No. 1, pp. 181-193, January 2006.
- [10] H. Zhu and M.C. Zhou, "Role-based collaboration and its kernel mechanisms," IEEE Transactions on Systems, MAN, and Cybernetics —Part C: Applications and Reviews, vol. 36, no. 4, pp. 578-589, July 2006.
- [11] M. VanHilst and D. Notkin, "Using role components in implement collaboration-based designs," In Proceedings of the 11th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications (OOPSLA '96). ACM, New York, NY, USA, vol. 31, no. 10, pp.359-369, October 1996.
- [12] H. Zhu, "A Role-based conflict resolution method for a collaborative system," IEEE International Conference on Systems, Man and Cybernetics, vol. 5, pp. 4135-4140, October 2003.
- [13] J. Y. Bang and et al., "CoDesign: a highly extensible collaborative software modeling framework," In Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - (ICSE '10), ACM, New York, NY, USA, vol. 2, pp. 243-246, 2010.
- [14] S. Herrmann, C. Hundt, and M. Mosconi, "OT/J language definition v1.3," Technical Universität Berlin. Object Teams home-page: <http://www.objectteams.org>, (2009), last visited: March 2012.
- [15] G. Kiczales and et al., "An overview of AspectJ," In Proceedings of the 15th European Conference on Object-Oriented Programming (ECOOP '01), Springer-Verlag, London, UK, pp. 327-353, 2001.
- [16] Object Management Group (OMG), CORBA v3.1 , Release Date: January 2008, Home-page: <http://www.omg.org/spec/CORBA/3.1/>, last visited: March 2012.
- [17] Sun Microsystems. "Java Remote Method Invocation Specification v1.5.0," Sun Microsystems Inc., Santa Clara, California,U.S.A. 2004.
- [18] R. Chitchyan and I. Sommerville, "Comparing dynamic AO systems," In Proceedings of the AOSD'04 - dynamic aspects workshop. Lancaster, UK, pp. 23-36, 2004.
- [19] A. Rocznik, J. Melhem, P. L'evy and A. El-Saddik, "Design of distributed collaborative application through service aggregation," In Proceedings of the 10th IEEE international symposium on Distributed Simulation and Real-Time Applications (DS-RT '06). IEEE Computer Society, Washington, DC, USA, pp.165-174, 2006.
- [20] N. Boulila, "Group support for distributed collaborative concurrent software modeling," In Proceedings of the 19th IEEE international conference on Automated software engineering (ASE '04). IEEE Computer Society, Washington, DC, USA, pp. 422-425, 2004.

Abdullah O. Al-Zaghameem (Ph.D) a lecturer in Tafila Technical University (Tafila, Jordan) since September, 1999. He was awarded the doctoral degree of Computer Science in Software Engineering from Technical University of Berlin (Germany) in September, 2012. His current research interests include: distributed AOP, collaboration-based computing, and Information Systems Design.

Mohammad Alfraheed (Ph.D) a lecturer in Tafila Technical University (Tafila, Jordan) since September, 2012. He has the doctoral degree of Computer Science in Computer Vision from RWTH-Aachen University (Germany) in July, 2012. His current research interests include: Digital Image Processing, Computer Vision, Robotics and Software Engineering.

Exit Chart for Iterative Decoding of Product and Concatenated Block Codes

Abderrazak Farchane¹, Mostafa Belkasmi² and Ahmed Azouaoui³

¹University Mohamed V SOUSSI, ENSIAS
Rabat, 10000, MOROCCO

²University Mohamed V SOUSSI, ENSIAS
Rabat, 10000, MOROCCO

³University Mohamed V SOUSSI, ENSIAS
Rabat, 10000, MOROCCO

Abstract

In this paper, we study the convergence behaviour of a new SISO decoder that we have proposed recently in [1]. The exchange of extrinsic information is used to trace the decoding trajectory in the extrinsic information transfer Chart. That allows foreseeing the turbo cliff position. The influence of concatenated scheme as well as different constituent codes on the convergence behaviour is investigated for product and generalized serially and parallel concatenated codes. Simulation result shows that the thresholds obtained by EXIT chart and BER curves are almost the same.

Keywords: EXIT chart, SISO decoder, iterative decoding, product codes, BCH codes, GSCB codes, GPCB codes.

1. Introduction

Stephan Ten Brink [2, 3] introduced a useful tool to analyse the convergence behaviour of iterative decoders. This tool is known as EXIT chart (The extrinsic information transfer chart). It allows us analysing the behaviour of iterative decoders that exchange soft information. This process of information exchange is described by a characteristic transfer diagram. The later shows the evolution of the extrinsic mutual information as a function of the a priori mutual information. The EXIT allows determining the $\frac{E_b}{N_0}$ beyond of that we can have a correct decoding. The EXIT chart can be used to design good codes that converge rapidly. It has been developed originally to analyse turbo and turbo-like codes based on convolutional codes. Subsequently, EXIT chart has been

extended to LDPC codes. Nevertheless, this technique is not directly applicable to Chase-Pyndiah's iterative decoder [4].

In [1], we have developed a new iterative decoder. It works for product and generalised concatenated block codes. This decoder uses a discrete parameter α . The later will be substituted by another continuous one. It is calculated using interpolation. The continuous parameter allows us applying the EXIT chart technique in order to study the convergence behaviour of product and generalised concatenated block codes. The simulation result shows that the thresholds obtained by the EXIT chart and BER chart are almost the same.

The remaining of this paper is organised as follows: the section two describes our decoder. We present the principle of EXIT chart in the section three. In section four, we introduce the parameter α for EXIT chart. Section five present the EXIT chart of our decoder. Finally, we conclude this paper in section six.

2. Description of our decoder

The decoding of product codes is done by decoding the rows, then the columns of the code matrix. Like turbo codes, it is possible to decode product codes using an iterative process. A reliability must be associated to each symbol. We consider a transmission that use BPSK modulation coded by a block code, with code rate $\frac{k_i}{n_i}$ ($i=1$ or 2). The input of the decoder, when the channel is perturbed by a white Gaussian noise, is equal to $Y = X + n$, where Y is the observed vector, The binary random variable denotes the transmitted bits with realizations $x \in \{\pm 1\}$; for brevity of notation, we will not distinguish between X and x in the following (only where

needed for clarification) and n is the white noise whose components n_j has zero average and variance σ^2 .

The reliability of component using the log-likelihood ratio (LLR) of the received sequence is defined by:

$$r_j = \ln \left(\frac{P_r [x = +1/y_j]}{P_r [x = -1/y_j]} \right) = \frac{2}{\sigma^2} y_j \quad (1)$$

Decoding of rows (or columns) is realized using a list decoding algorithm that lets us to determine the most likelihood codewords. Then, among those codewords, we select the closest codeword to the received sequence R in term of euclidean distance, where $R = (r_1, r_j, \dots, r_n)$ is the reliability of the received sequence.

The decoder affects a weighting to each component of the decided codeword d_j , in order to measure the reliability of each decision. This reliability is evaluated by the logarithm of the likelihood ratio associated to a decision d_j at the output of the decoder and is defined by

$$\Lambda_j = \ln \left(\frac{P_r [d_j = +1/R]}{P_r [d_j = -1/R]} \right) \quad (2)$$

The sign of Λ_j gives the decision d_j and the absolute value of Λ_j is the measure of the reliability of this decision. When the signal to-noise ratio is sufficiently big and the noise is Gaussian, the LLR of (2) can be simplified to this form:

$$\Lambda_j = \frac{1}{2\sigma^2} \left(\|R - C^{\min(-1)}\|^2 - \|R - C^{\min(+1)}\|^2 \right) \quad (3)$$

where the two codewords $C^{\min(+1)}$ and $C^{\min(-1)}$ having the minimum distance from R and belonging respectively to S_j^{+1} and S_j^{-1} . By introducing the components of the vector R and if we suppose that σ is constant, we can normalize Λ_j with respect to the constant $\frac{2}{\sigma^2}$. We can write the LLR in following form [4]:

$$\begin{aligned} \hat{\Lambda}_j &= \frac{\sigma^2}{2} \Lambda_j \\ &= \frac{1}{4} \left(\|R - C^{\min(-1)}\|^2 - \|R - C^{\min(+1)}\|^2 \right) \quad (4) \\ &= r_j + e_j \end{aligned}$$

The LLR of a bit is equal to the sum of the reliability of simple r_j in the input of the decoder and a quantity e_j is independent to the reliability of simple r_j . The quantity e_j is analogue to the extrinsic information for the convolutional turbo codes.

In order to determinate the simplified expression of the LLR of a bit in the output of the decoder, it is necessary to determinate the two codewords $C^{\min(+1)}$ and $C^{\min(-1)}$ having the minimum distance from R and having an opposite sign in position j . For this, we use a list decoding algorithm (like Chase algorithm [5]). It allows us to determine a sub set of codewords among which we can find the searched two codewords.

Sometimes we can not find the two codewords in sub set determined by the list decoding (Chase algorithm for example). This means that all codewords have the same decision on the j^{th} element, d_j of the vector D . They vote for the same

candidate. In this case the decision confirms the input decoder. Consequently, the reliability of the decision must be increased while the sign of decision, d_j , is given by the decoder. We propose a formula that can allows computing the reliability of the decision by taking into account the reliability of the decoder input, the sign of the decision. The $\hat{\Lambda}_j$ of the j^{th} element of the decision is given by the following formula:

$$\hat{\Lambda}_j = \frac{1}{2} \sigma_R d_j + |r_j| d_j \quad (5)$$

where σ_R represents the standard deviation of the decoder input.

The extrinsic information $e_j(p)$ is exploited to modify the input of the next decoder. The reliability of the p^{th} decoder is given by the following formula:

$$r_j(p) = r_j + \alpha(p) e_j(p) \quad (6)$$

The parameter $\alpha(p)$ is used to control the reliability of the extrinsic information that is unreliable in the first few iterations. However, it became reliable with the iterations.

Using the equations 2 and 4, the extrinsic information e_j is evaluated in the p^{th} decoder by the formula:

$$\begin{aligned} e_j(p) &= \frac{1}{4} \left(\|R(p) - C^c(p)\|^2 - \|R(p) - D(p)\|^2 \right) \\ &\quad \times d_j(p) - r_j(p) \quad (7) \end{aligned}$$

where $r_j(p)$ represents the j^{th} component of the vector $R(p)$ at the input of the p^{th} decoder. $D(p)$ is the decided codeword at the output of the p^{th} decoder. $C^c(p)$ is the competitor codeword such that the symbol at the j^{th} position is opposite to d_j ($c_j^c = -d_j$).

The figure 1 shows the decoding scheme of the product and generalized concatenated codes.

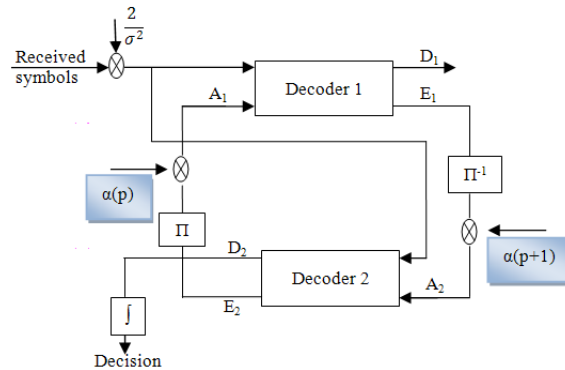


Figure 1: Product and generalized concatenated codes scheme [1]

3. Principle of EXIT chart

EXIT chart is a semi-analytical tools originally developed by S.T. Brink. It allows us to visualise the decoding trajectory

of iterative decoders. Its transfer characteristic curves are obtained by plotting the mutual information between the input and the output of the constituent decoders. The iterative decoding of concatenated codes is done by two soft constituent decoders. The mutual information at the output of the first decoder is re-injected in the input of the second decoder. By the same manner the output of the second decoder is connected to the input of the first. It is possible to obtain independently the transfer characteristic of the two decoders. Finally, by combining the results in the same figure we obtain the EXIT chart of the iterative decoder. This technique has the advantage that it doesn't demand time as many as the BER chart. Indeed, by simulation, it has been demonstrated that the a priori returned by the decoder is Gaussian distribution. Consequently, it is unnecessary to consider the whole scheme but we take only one constituent decoder whose a priori is generated by using a Gaussian distribution.

3.1. LLR of the AWGN channel

It is suitable to use the notations of the figure 1. Let A_i denote the a priori information, where $i = 1, \text{ or } 2$, E_i denote the extrinsic information, D_i denote the output of the decoder, and the soft input is denoted by y . For the AWGN channel, the input and output variables are considered to be random variables that are related by

$$Y = X + n \tag{8}$$

where X is a random variable representing the bits x of the message. The bit x can take two values $x = +1$ or $x = -1$. Let Y be a random variable representing the channel observation, n be a random variable denoting the noise of variance $\sigma^2 = \frac{N_0}{2}$. The logarithm likely-hood ratio, LLR, of the channel is denoted by the following equation:

$$L_c = \log \left(\frac{P(y|x = +1)}{P(y|x = -1)} \right) = \frac{2}{\sigma^2} y \tag{9}$$

where

$$P(y|x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}} \tag{10}$$

the expression (9) allows to determine the random variable Y that represent the received sequence:

$$L_c = R = \frac{2}{\sigma^2} y = \frac{2}{\sigma^2} (x + n) \tag{11}$$

This can also be written as

$$R = \mu_R x + n_R \tag{12}$$

where

$$\mu_R = \frac{2}{\sigma^2} \tag{13}$$

and n_R is a random variable of zero mean and variance:

$$\sigma_R^2 = \frac{4}{\sigma^2} \tag{14}$$

The mean value and the variance of this random variable are related by

$$\mu_R = \frac{\sigma_R^2}{2} \tag{15}$$

This relationship will be useful in the construction of the EXIT chart.

We have assumed that the observation and the a priori have a Gaussian distribution. Consequently, even if the processing done by the decoder isn't linear it's natural to consider that probability of the extrinsic at the output of the constituent decoder is also Gaussian.

3.2. EXIT chart model and mutual information for consistent Gaussian channel

It is useless to constitute the whole decoding scheme. But, we consider the functioning of one constituent decoder or each constituent decoder solely. The a priori is artificially generated. The figure 2 shows the model used to obtain the EXIT chart.

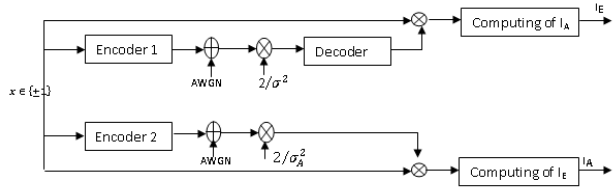


Figure 2: Calculating EXIT chart model.

The prediction of convergence behaviour of iterative decoders can be done by determining the relation between the input and the output of the constituent decoders. For this we use some observation obtained by simulation [6, 7].

1. The a priori A remains uncorrelated with the channel observation for several iterations and wide size interleavers.
2. The probability density function of the extrinsic can be approached by a Gaussian distribution.

These observations are done for the PCC codes in [6], and we have verified it, in section 5, for the GPCB codes using the proposed decoder.

These two observations allow modelling the a priori by:

$$A = \mu_A x + n_A \tag{16}$$

where n_A is a Gaussian random variable of zero mean and variance σ_A^2 . The mean μ_A verifies the following relation:

$$\mu_A = \frac{\sigma_A^2}{2} \tag{17}$$

The conditional probability density function of the a priori A is given by:

$$P_A(\xi|X = x) = \frac{1}{\sqrt{2\pi}\sigma_A} e^{-\frac{(\xi - \frac{\sigma_A^2}{2}x)^2}{2\sigma_A^2}} \tag{18}$$

The a priori mutual information $I_A = I(X; A)$, between the variable A and the variable X that represents the systematic

bits is used to quantifies the a priori information. The mutual information can be calculated as follow [8, 9]:

$$I_A = \frac{1}{2} \sum_{x=-1,+1} \int_{-\infty}^{+\infty} P_A(\xi|X=x) \times \log_2 \left(\frac{2P_A(\xi|X=x)}{P_A(\xi|X=-1) + P_A(\xi|X=+1)} \right) d\xi \quad (19)$$

where $0 \leq I_A \leq 1$

The expression (19) can be combined with (18) in order to develop the expression of the mutual information as follows:

$$I_A = 1 - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_A} e^{-\frac{(\xi - \frac{\sigma_A^2}{2})^2}{2\sigma_A^2}} \log_2(1 + e^{-\xi}) d\xi \quad (20)$$

For a given value of σ_A , we can compute numerically or by simulation the a priori mutual information, I_A . Figure 3 represents the curve of mutual information I_A as a function of σ_A by using simulation. We can approximate (19) by [10]:

$$I_A \approx 1 - \frac{1}{N} \sum_{i=1}^{i=N} \log_2(1 + e^{-xL_{A_i}}) \quad (21)$$

where N is the size of the transmitted message.

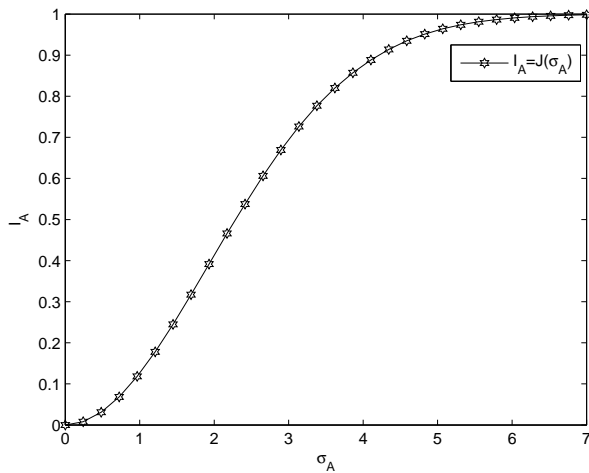


Figure 3: The a priori mutual information, I_A , as a function of σ_A .

Let $J(\sigma_A)$ be the abbreviation of $I_A(\sigma_A)$, we have the following properties.

- $0 \leq J(\sigma_A) \leq 1$
- $\lim_{\sigma_A \rightarrow 0} J(\sigma_A) = 0$
- $\lim_{\sigma_A \rightarrow +\infty} J(\sigma_A) = 1$

the function $J(\sigma_A)$ is monotonous. So $J(\sigma_A)$ is reversible. Let J^{-1} be the inverse function defined by:

$$\sigma_A = J^{-1}(I_A) \quad (22)$$

Mutual information is used to quantify extrinsic information $I_E = I(X; E)$.

$$I_E = \frac{1}{2} \sum_{x=-1,+1} \int_{-\infty}^{+\infty} P_E(\xi|X=x) \times \log_2 \left(\frac{2P_E(\xi|X=x)}{P_E(\xi|X=-1) + P_E(\xi|X=+1)} \right) d\xi \quad (23)$$

4. The Parameter α for EXIT chart

In the decoding scheme of product codes [4, 11], the a priori information is scaled by a parameter α . The value of the later parameter increases with the iterations. During the first iterations, α limits the effect of the extrinsic information because it is less reliable. Nevertheless, α increases in proportion as the extrinsic information gets improved. Therefore, the values of α have to be harmonized with extrinsic information exchanged between the two constituent decoders.

The EXIT Chart allows us to study the convergence of iterative decoders, by studying only one constituent decoder or each constituent decoder separately. This study is based on the exchange of the extrinsic information between constituent decoders.

In the works [2, 3, 6], the authors study the product codes and concatenated codes. Their decoding scheme consists to inject the extrinsic information without scaling it at the input of the next decoder. However, our decoding scheme [1] and also other schemes in [4, 11] consist to scale extrinsic information by a scaling factor α . It is difficult to study the convergence of product and generalised concatenated block codes using EXIT Chart technique because we haven't the notion of iterations there. Nevertheless, we change the value of the a priori mutual information and follow the evolution of the extrinsic mutual information. Both the a priori and extrinsic information must be scaled by the factor α . So, how can we choose the value of the parameter α that goes with a priori and extrinsic information?

In order to resolve this problem, we have to find a function allowing to calculate the parameter α whatever the a priori information. One solution is to interpolate the discrete value of α by a continual function. This needs to express this parameter as a function of another, like the standard deviation or the mutual information. Once we have a couple of points, the standard deviation and it's corresponding value of α we can determine the interpolation function that interpolates these points. Finally, the interpolation function permits to calculate the value of the parameter α that goes well with the a priori and extrinsic information.

To interpolate the parameter α , we fix the code, for example $BCH(63, 51)^2$, and we take the SNR that gives $BER = 10^{-5}$. The value of the standard deviation of the extrinsic information is used to interpolate α are shown in table 1 : We use the interpolation by pieces to interpolate α . The curve of the figure 4 represents the function that interpolates α as a function of the standard deviation σ_E of extrinsic information.

α	0.0	0.13	0.15	0.18	0.2	0.25
σ_E	0.0	3.34	3.61	3.75	3.9	4.04
α	0.3	0.35	0.4	0.45	0.5	0.55
σ_E	4.3	4.65	5.13	5.72	6.34	6.95
α	0.6	0.65	0.7	0.72	0.75	0.77
σ_E	7.59	8.3	9.12	10.08	10.91	11.82
α	0.8	0.82	0.85	0.87		
σ_E	12.70	13.75	14.83	16.14		

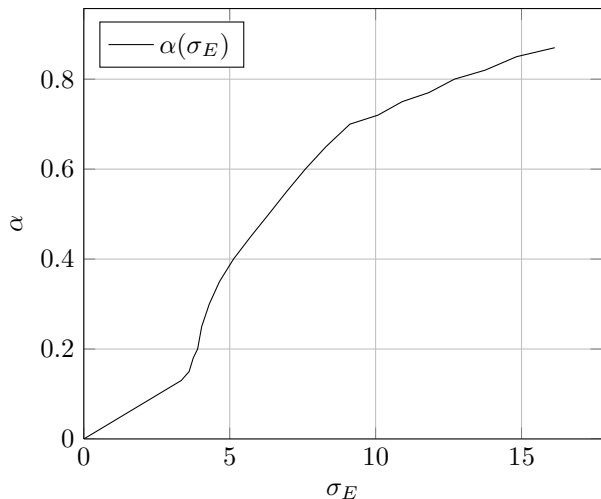


Figure 4: Parameter α as a function of σ_E of extrinsic information for the product code $BCH(63, 51)^2$.

To verify the effectiveness of this way (interpolation), we evaluate the performance of the code $BCH(63, 51)^2$ using interpolated α . The figure 5 shows the performance of the product code $BCH(63, 51)^2$. According to this figure and the result obtained in [1] we observe that the discrete and interpolated parameters are equivalent. Thus, the EXIT chart for product and generalised concatenated block codes will be done by using interpolated α .

5. EXIT chart for our decoder

Monte Carlo simulation allows us to affirm that the values of the a priori information A are independent and uncorrelated with the observations. On the other hand, the probability density function, or more exactly, the histogram of the values of the extrinsic information is Gaussian distribution. It is also true that these values become the a priori values for the next iteration, so that a similar conclusion can be stated for the probability density function of the a priori information. That is to say, the extrinsic and the a priori can be approached by a Gaussian distribution.

The figure 6 presents the non-normalized histogram of the extrinsic information generated by the decoder in [1] for the code $GPCB-BCH(141, 113)$ with $M = 500$ [12]. A zero

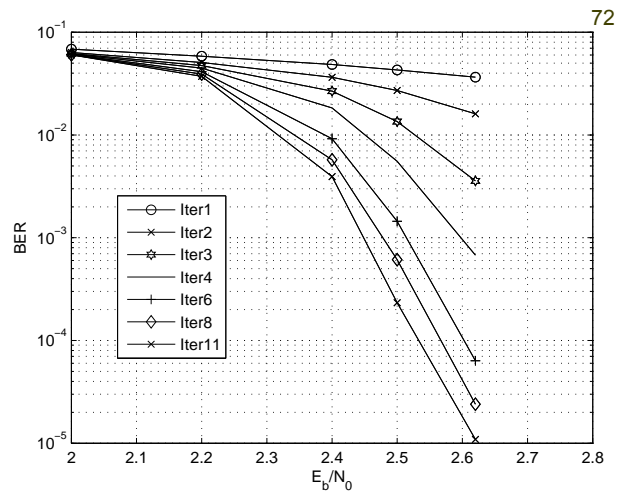


Figure 5: Performance of the decoder in figure 1 for the product code $BCH(63, 51)^2$ over AWGN channel, with interpolated α .

codeword is transmitted. The curves are plotted for different SNR: $\frac{E_b}{N_0} = 3, 4$ et $5dB$. Increasing of the SNR renders the histogram of the decoder of figure 1 moves towards the left. According to this figure, the histogram can be approached by Gaussian distribution. To make sure that the distribution of

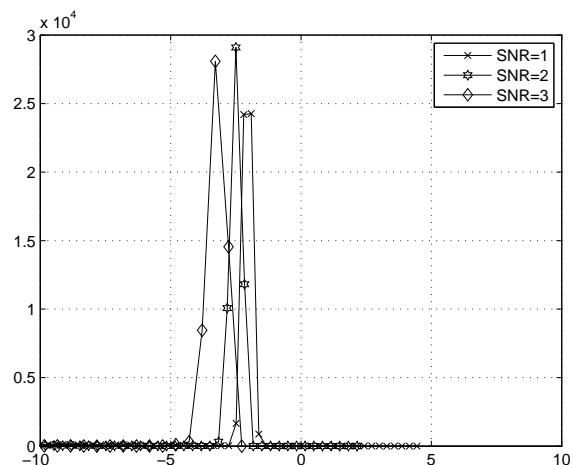


Figure 6: Non-normalized histogram of extrinsic information E at the output of the decoder for $GPCB-BCH(141, 113)$ code, with interleaver size equal to 500×113 .

the a priori and extrinsic information are Gaussian distribution, we feed the decoder of the figure 2 by a message of length 500×113 . This message will be coded using the code $GPCB-BCH(141, 113)$ with $M = 500$. Then we feed the decoder by a priori information of variance $\sigma_A = 2$ et 4.5 . This information is generated according to the equation (16) and (17), with the parameter $\frac{E_b}{N_0} = 1.5dB$. The histograms of the a priori and extrinsic are represented by the figure 7. According to this figure, we remark that the increasing of σ_A lets the histograms move toward the right or the left, depending on the transmitted messages (null vector or full 1 vector).

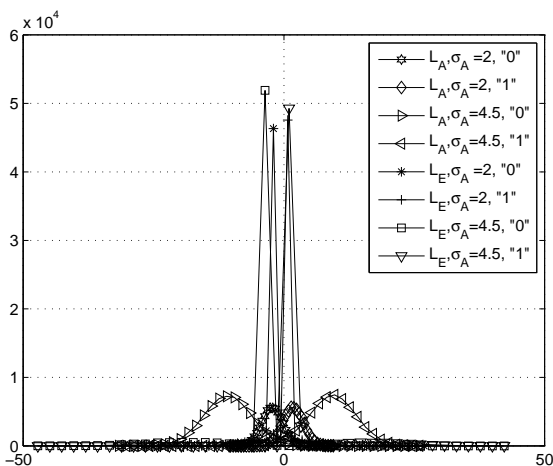


Figure 7: Non-normalized histogram of the a priori information applied to the decoder in figure 1, and the resultant non-normalized histogram of the extrinsic information for the code $GPCB - BCH(141, 113)$ with $M = 500$, over AWGN channel.

5.1. Convergence behaviour analysis of product codes

The EXIT Chart describes the relation between the mutual information of the a priori, I_A , and extrinsic, I_E . Both the a priori and the extrinsic information are measured by using the mutual information between these quantities and the information in the systematic or message bits. The transfer diagram can be obtained by computing, for a given value of I_A and a fixed value of the parameter $\frac{E_b}{N_0}$, the corresponding value of I_E . This computing assumes that the a priori have been generated using the relations (16) and (17), where the probability density function is described by (18). The a priori information A is applied to the constituent decoder, of the figure 2, with a codeword affected by a Gaussian noise according to the value of $\frac{E_b}{N_0}$. The decoder generates extrinsic information E that can be quantified by I_E . The value of the later variable can be obtained by the Monte Carlo method using the relation(23).

The figure 8 shows the transfer function $T(I_A)$, for a set of values of the parameter $\frac{E_b}{N_0}$, for the code $BCH(63, 51)^2$ of rate 0.65. The a priori I_A represents the abscissa axes, and the extrinsic I_E represents the ordinate axes. According to this figure, we remark that the increasing of the value of $\frac{E_b}{N_0}$ involves the increasing of the extrinsic mutual information I_E . The parameter α used to obtain the transfer characteristics is interpolated α . The interpolation of the parameter α , by a continuous function, allows to choose the appropriate value that can goes with the a priori.

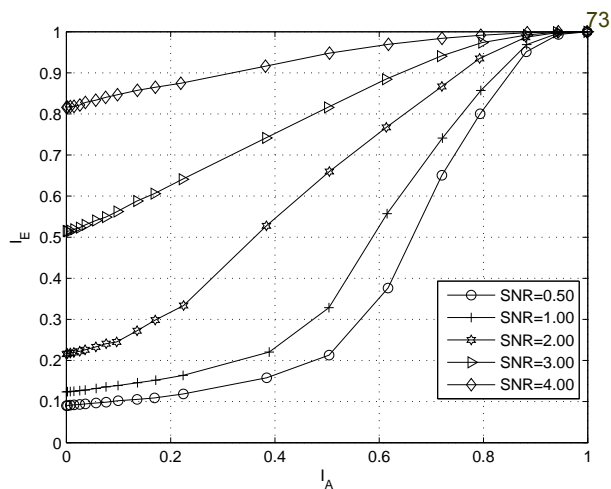


Figure 8: Transfer characteristics of the product code $BCH(63, 51)^2$, with code length 63×63

5.2. Trajectory of iterative decoding

The trajectory of iterative decoding is obtained as follows: Let n denotes the index of iterations and E_b/N_0 a fixed signal to noise ratio. At $n = 0$ the decoder starts with a zero a priori, $I_{A1,0} = 0$. At the iteration n the extrinsic of the first decoder $I_{E1,n} = T_1(I_{A1,n})$ is submitted to the second decoder as a priori $I_{A2,n} = I_{E1,n}$. The extrinsic information of the second decoder $I_{E2,n} = T_2(I_{A2,n})$ is re-injected in the first decoder as a priori $I_{A1,n+1} = I_{E2,n}$. The iterations are stopped when $I_{E2,n+1} = I_{E2,n}$, this corresponds to the intersection of the two EXIT charts. The process of iterative decoding designs theoretically a trajectory by projection over the transfer characteristic of the decoder.

To describe the nature of iterative decoding, the characteristic curves of the two iterative decoders must be plotted in the same figure. However, the axes of the transfer characteristic of the second decoder are swapped. The exchange of the mutual information can be seen as a decoding trajectory. The later can be obtained by designing the zigzag over the EXIT chart, it must coincides with the one obtained by simulation.

The figure 9 depicts the EXIT chart of the product code $BCH(63, 51)^2$ for the SNR 0.5, 1.2, 2.0, 2.3, 2.5, and 3.0 dB. Note that in the graphical representations the decoder characteristics are only plotted up to their first intersection. An opening for the trajectory at 2.3dB can clearly be seen, that corresponds to the turbo cliff position in BER chart of the figure 5.

The main contribution of the EXIT chart to the analysis of iterative decoding is the advantage that only simulation on individual decoder is needed to obtain the desired transfer characteristics. No BER simulation of the iterative decoding scheme itself is required. Moreover, for turbo code with the same constituent decoder, transfer characteristics of only one constituent decoder are sufficient to predict the performance of iterative decoding. This further speed up the evaluation of new concatenated codes.

The figure 10 represents the iterative decoding trajectory

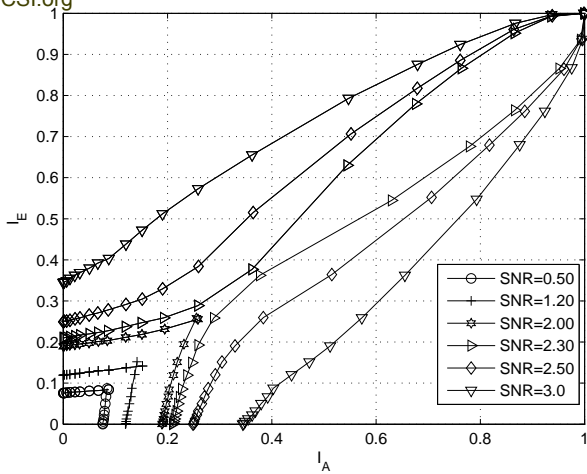


Figure 9: EXIT Chart of our decoder for the product code $BCH(63, 51)^2$, with code length 63×63 .

drawn as a zigzag over EXIT chart at the $SNR = 2.5dB$. After three passes through the decoder, increasing correlations of extrinsic information starts to show up and lets the trajectory deviate from its expected zigzag-path.

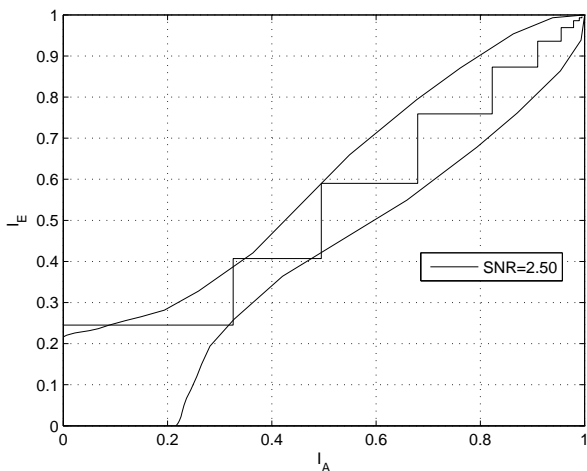


Figure 10: Simulated trajectory of iterative decoding of the product code $BCH(63, 51, 5)^2$ at $SNR = 2.5dB$, with code length 63×63 .

The figure 11 shows the EXIT chart of the product code $BCH(127, 113)^2$ for the SNR 1.0, 1.5, 2.2, 2.5, 2.9, 3.0, 3.5, 4.0 and 4.5 dB. For $SNR = 3.0dB$, the decoding trajectory enters in tunnel region close to the bisector. The threshold obtained by the EXIT chart is 0.1dB away from the starting point of the waterfall region of the BER curves as shown in the figure 12. The difference between the estimated threshold obtained by EXIT chart and that obtained by BER chart is due to the fact that the size of codeword is finite, and due to the fact of computing α using interpolation.

The figure 13 represents the transfer diagram of the product code $BCH(255, 247)^2$ for the SNR 1, 1.5, 3.0, 4.3, 4.6 and

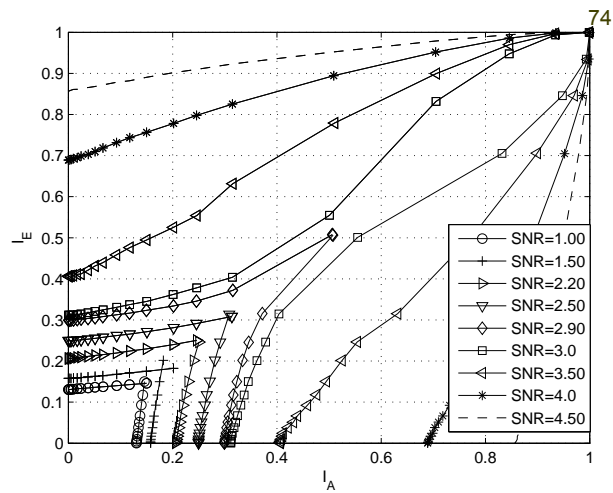


Figure 11: EXIT Chart of our decoder for the product code $BCH(127, 113)^2$, with code length 127×127 .

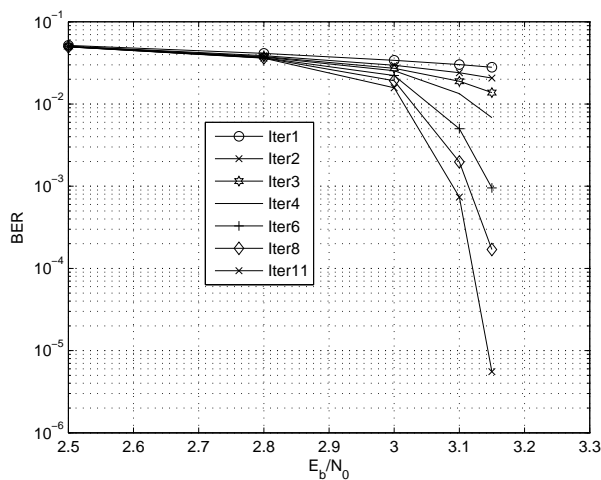


Figure 12: Performance of our decoder for the product code $BCH(127, 113)^2$.

5.0 dB. For $SNR = 4.3dB$, the decoding trajectory enters in tunnel region close to the bisector. The threshold obtained by the EXIT chart is 0.2dB away from the starting point of the waterfall region of the BER curves of the figure 14. The difference between the estimated threshold obtained by EXIT and that obtained by BER curves can be explained by the fact that the size of codeword is finite and approximation done by interpolation.

The figure 15 displays the transfer diagram of the product code $BCH(511, 493)^2$ for the SNR 2.0, 3.5, 4.3, 4.5, and 4.7 dB. For $SNR = 4.2dB$, the decoding trajectory enters in tunnel region close to the bisector. The threshold obtained by the EXIT chart is 0.2dB away from the starting point of the waterfall region of the BER curves of the figure 16. The difference between the estimated threshold obtained by EXIT and that obtained by BER curves can be explained

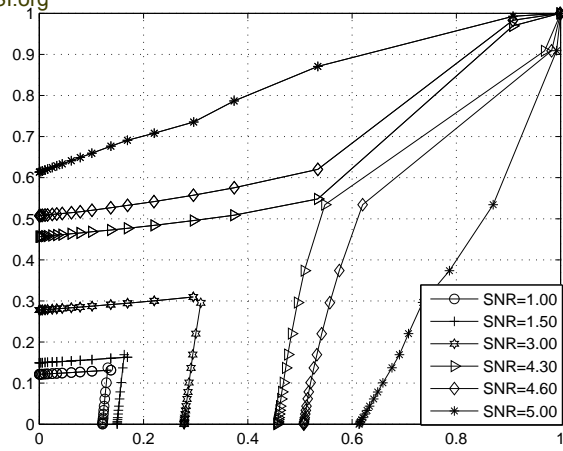


Figure 13: EXIT Chart of our decoder for the product code $BCH(255, 247)^2$, with code length 255×255 .

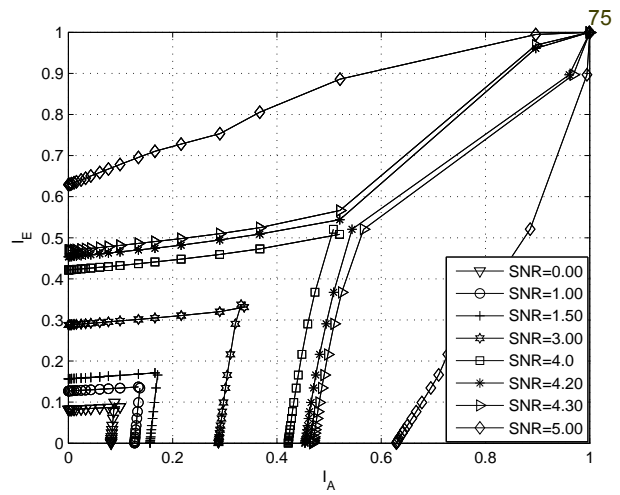


Figure 15: EXIT Chart of our decoder for the product code $BCH(511, 493)^2$, with code length 511×511 .

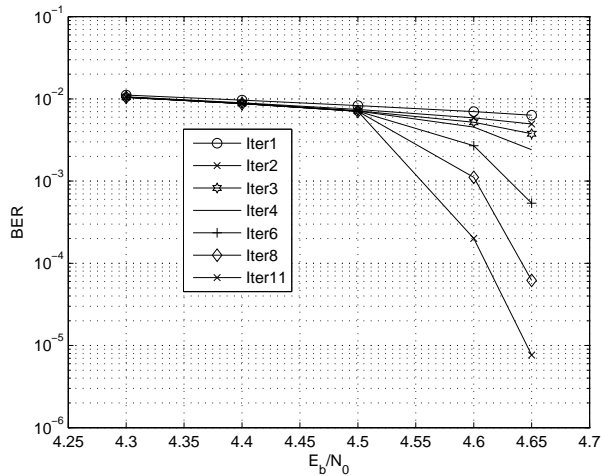


Figure 14: Performance of our decoder for the product code $BCH(255, 247)^2$

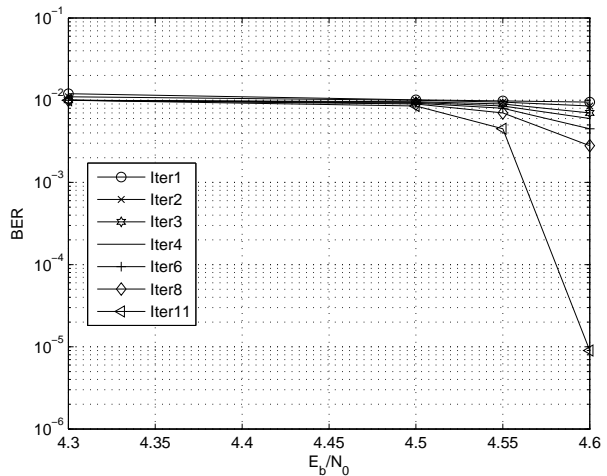


Figure 16: Performance of our decoder for the product code $BCH(511, 493)^2$

by the fact that the size of codeword is finite and approximation done by interpolation.

5.3. Convergence behaviour analysis of GPCB codes

The figure 17 shows the transfer diagram of the parallel concatenated code $GPCB - BCH(75, 51)$ for the SNR 1.0, 1.5, 2.0, 2.45 and 3.0 dB. For $SNR = 2.45dB$, the decoding trajectory enters in tunnel region close to the bisector. The threshold obtained by the EXIT chart is 0.15dB away from the starting point of the waterfall region of the BER curves of the figure 18. The difference between the estimated threshold obtained by EXIT and that obtained by BER curves can be explained by the fact that the size of codeword is finite and approximation done by interpolation.

The figure 19 displays the transfer diagram of the parallel

concatenated code $GPCB - BCH(75, 51)$ for the SNR 1.0, 2.0, 2.8, 2.9, 3.0, 3.5 and 4.0 dB. For $SNR = 2.9dB$, the decoding trajectory enters in tunnel region close to the bisector. The threshold obtained by the EXIT chart coincides with the threshold of the waterfall region of BER chart of the figure 20.

The figure 21 represents the trajectory of the iterative decoding designed over the EXIT chart of the $GPCB - BCH(75, 51)$, for the $SNR = 2.6dB$. The trajectory has just managed to sneak through the bottleneck. After three passes through the decoder, increasing correlations of extrinsic information starts to show up and lets the trajectory deviate from its expected zigzag-path.

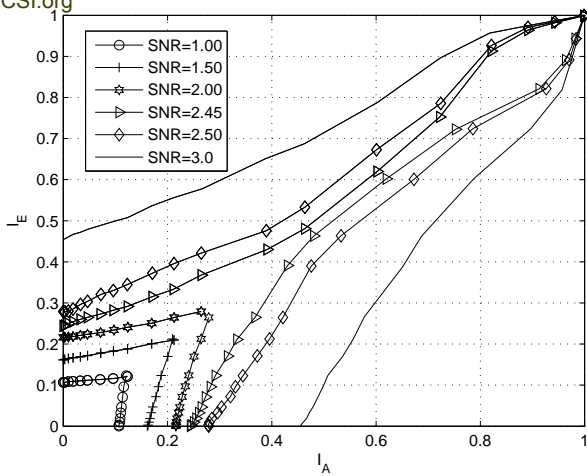


Figure 17: EXIT Chart of our decoder for the parallel concatenated codes $GPCB - BCH(75, 51)$, with $M=100$ over AWGN channel

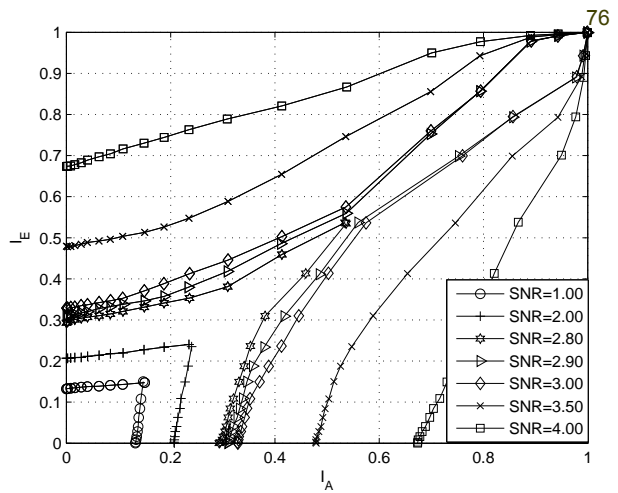


Figure 19: EXIT Chart of our decoder for the parallel concatenated codes $GPCB - BCH(141, 113)$, with $M=100$, over AWGN channel

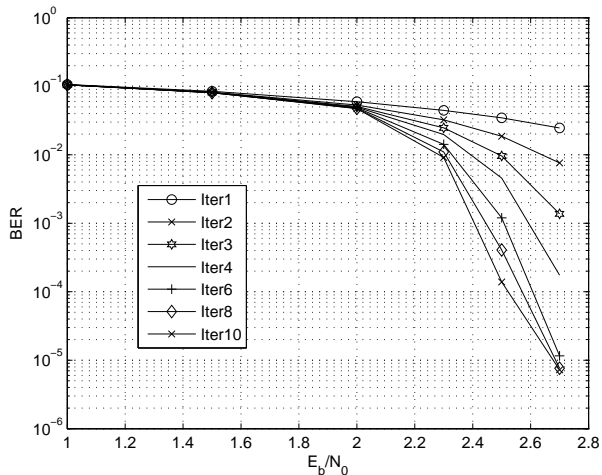


Figure 18: Performance of our decoder for the parallel concatenated code $GPCB - BCH(75, 51)$, $M=100$.

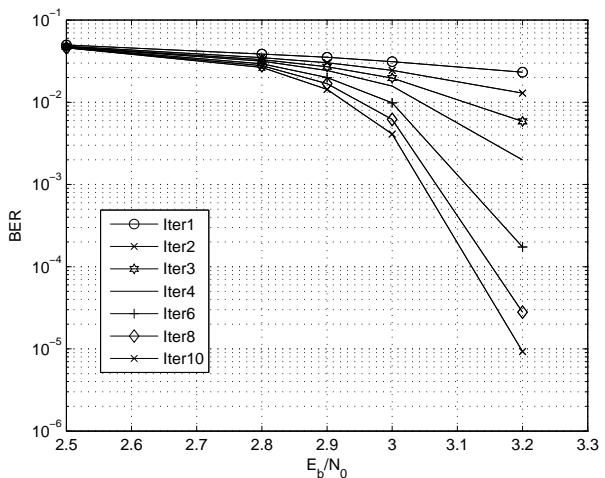


Figure 20: Performance of our decoder for the parallel concatenated code $GPCB - BCH(141, 113)$, with $M=100$.

5.4. Convergence behaviour analysis of GSCB codes

We have seen that for symmetric product codes and symmetric generalized parallel concatenated code the studying of unique decoder is sufficient to foresee the behaviour of the iterative decoding. Nevertheless, in the case when we have non symmetric elementary decoders the EXIT chart necessitates studying both constituent decoders independently. The figure 22 shows the transfer characteristic of the generalized serially concatenated code $GSCB-BCH(63, 39)$ [13, 1]. According to this figure, the threshold is achieved at $SNR = 2.3dB$.

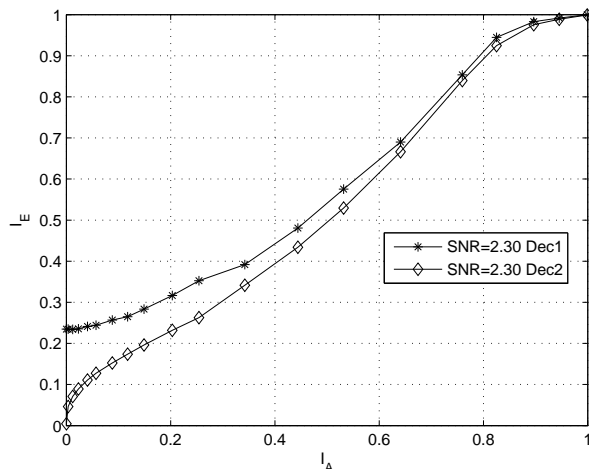


Figure 22: EXIT Chart of our decoder for the serially concatenated code $GSCB - BCH(63, 39)$, with $M=100$

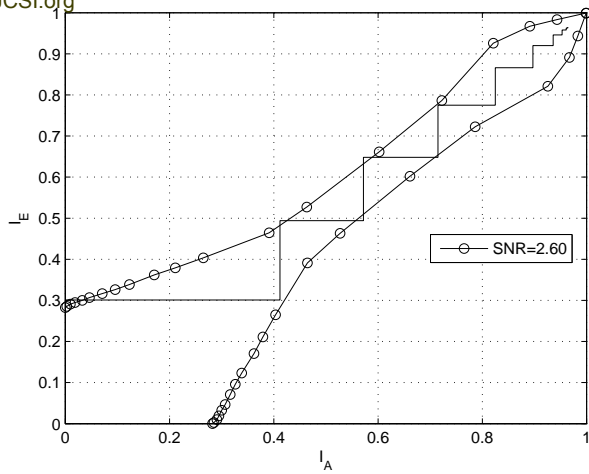


Figure 21: EXIT chart at $SNR = 2.6dB$ of parallel concatenated code $GPCB - BCH(75, 51)$, with $M=100$

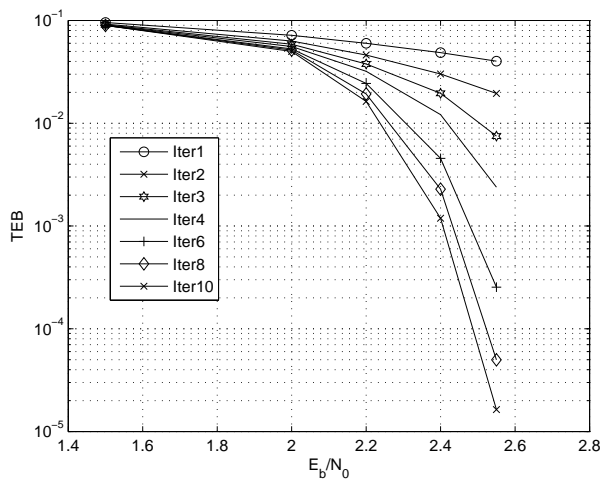


Figure 23: Performance of our decoder for the serially concatenated code $GSCB - BCH(63, 39)$, with $M = 100$.

6. Conclusion

In this work, we have analysed the convergence behaviour of iterative decoder of product and generalized concatenated block codes using the technique called EXIT chart. According to the results obtained by the BER and EXIT charts, we remark that the thresholds obtained by the two techniques are almost the same. The difference between the thresholds of the EXIT and BER charts may be explained by the fact that the size of codeword is finite and approximation done by the interpolation of the parameter α . Finally, EXIT chart is an alternative technique for designing iterative decoders. This technique can be extended to other codes, like RS codes, either over AWGN or Rayleigh channel.

References

- [1] A. Farchane and M. Belkasmı, "New efficient decoder for product and concatenated block codes," *Journal of Telecommunication*, vol. 12, pp. 17–22, JANUARY 2012.
- [2] S. Ten Brink, "Convergence behaviour of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, October 2001.
- [3] S. T. Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, May 1999.
- [4] R. Pyndiah, A. Glavieux, A. Picart, and S. Jacq, "Near optimum decoding of product codes," *GLOBECOM94*, November 1994.
- [5] D. Chase, "Class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. Information theory*, vol. 13, pp. 170–182, Jan 1972.
- [6] S. T. Brink, "Iterative decoding trajectories of parallel concatenated codes," *Proceedings of 3rd IEEE/ITG Conference on Source and Channel Coding, Munich, Germany*, January 2000.
- [7] S. T. Brink, *Design of Concatenated Coding Schemes based on Iterative Decoding Convergence*. PhD thesis, Institute of Telecommunications, University of Stuttgart, 2001.
- [8] R. W. Hamming, "Coding and information theory," *Prentice Hall, New Jersey*, 1986.
- [9] N. F. Kiyani and J. H. Weber, "Exit chart analysis of iterative demodulation and decoding of MPSK constellations with signal space diversity," *Journal of Communications*, vol. 3, pp. 43–50, JULY 2008.
- [10] J. Hagenauer, "The exit chart- introduction to extrinsic information transfer in iterative processing," *EUSIPCO*, 2004.
- [11] A. Picart and R. Pyndiah, "Adapted iterative decoding of product codes," *Global Telecommunications Conference GlobCom*, 1999.
- [12] A. Farchane, M. Belkasmı, and S. Nouh, "Generalized parallel concatenated block codes based on BCH and RS codes: construction and Iterative decoding," *Journal of Telecommunication*, vol. 12, pp. 1–9, JANUARY 2012.
- [13] A. Farchane and M. Belkasmı, "Generalized serially concatenated codes: construction and iterative decoding," *International Journal of Mathematical and Computer Sciences*, 2010.

Abderrazak Farchane received his license in Computer Science and Engineering in June-2001 and Master in Computer Science and telecommunication from University of Mohammed V - Agdal, Rabat, Morocco in 2003. recently, he obtained his PhD in Computer Science and Engineering at ENSIAS, Rabat, Morocco. His areas of interest are Information and Coding Theory.

Protection of Personal Data in Information Systems

Bojken Shehu ¹, Shqiponja Ahmetaj ², Marin Aranitasi ³ and Aleksander Xhuvani⁴

¹ Computer Engineering Department, Faculty of Information Technology, Polytechnic University of Tirana
Tirana,Albania

² Computer Department, Faculty of Informatics, Vienna University of Technology
Vienna,Austria

³ Center for R&D, Faculty of Information Technology, Polytechnic University of Tirana
Tirana,Albania

⁴ Computer Engineering Department, Faculty of Information Technology, Polytechnic University of Tirana
Tirana,Albania

Abstract

The rapid development of computing technology has led to the emergence of the greater capacity to store personal data. The huge amount of information that governments and businesses collect from individuals have become a cause of concern. Personal data collection encroaches on the individual's right, particularly as it invades privacy or the right to control information about ourselves; there is a disclosure of private personal facts; and, the information can be used in such a way that a person is cast in a bad light or in a case of identity theft. The method of personal data collection, its use and security, hence, necessitate citizen protection. Through the years, laws that aim to protect personal privacy have enacted but they appear to be insufficient. This paper examines the concept of depersonalization as an alternative method for the representation and protection of personal information. It is based on the argument that the legal protection available is not sufficient to address privacy concerns.

Keywords: *Data Security, Depersonalization, Legal Protection, Personal Data.*

1. Introduction

The rapid development of computing technology has led to the emergence of the greater capacity to store personal data. These data are important and collected because of research. Governments collect them for the evaluation of its programs or for use databases such as in law enforcement and social security. Businesses are voracious consumers of these data because they are crucial in the success of several operations because they record habits, preferences, among other patterns of individual activities that help them develop market and sell products and services. Computers and the Internet paved the way for more efficient and fast methods to gather, store and

organize personal information. Since the 1970s the number of computer data banks or databases became staggering. There are the databases from Social Security Administration, Law and Order authorities (like the Federal Bureau of Investigation in the USA), Medical Information Bureau, state criminal justice systems, municipal data systems, credit card companies, telephone companies and recently Google, Facebook and so much more.

This paper will explore depersonalization as an effective method of personal data collection, where privacy is still protected. The description and analysis of depersonalization reveal a sound framework that can achieve a level of privacy protection in a manner that does not hinder the need of governments and businesses for statistical research. These data are important and collected for statistical purposes. Governments collect them for the evaluation of their programs or use them as databases in different areas. Businesses are voracious consumers of these data which are crucial in the success of several operations. They record habits, preferences and other patterns of individual activities that help them develop, market and sell products and services. Computers and the Internet paved the way for more efficient and fast methods to gather, store and organize personal information.

2. Legal Protection of Personal Data

The foundation of legal protection against indiscriminate collection of private data is a Congressional report that outlined four tenets of fair information practices, namely:

- Notice or the disclosure of the details of data gathering practices, policies and results to data subjects.
- Choice or the ability of the data subjects to exercise choices about how their personal data can be used.
- Access or the level of access provided to individuals on the gathered data about them.
- Security or the responsibility of data gatherers to provide adequate protection for the information collected [1] (Bidgoli 2004, page.98).

Based from these principles, a number of laws were enacted covering individual privacy across different sectors. For instance, the Gramm-Leach-Bliley Act protects personal banking information; the Fair Credit Reporting Act provides the framework for handling personal credit data. There are also laws that cover the collection and use of medical and health data, government records, children’s privacy, and so forth. Laws are also enacted in other countries such as the European Union Data Protection Directive, the OECD privacy guidelines adopted by countries such as Mexico, Australia, Japan and Czech Republic [3] (Conrad, Misener and Feldman 2012, page.401). But these laws and even some ethical guidelines [5] (Kluge 2000), no matter how specific and comprehensive, still fail to address privacy issues. Neubauer and Kolb [7] (2009) , for example, noted that approaches and methods for protecting privacy often do not comply with legal requirements or basic security requirements without suffering any penalty, (7). Szeto and Miri [10] (2007) revealed the same findings when they studied the Canadian experience. According to Hildebrandt and Gutwirth [4] (2008), this is because most statutes builds on traditional ways of thinking data, personal data and their abuse, without understanding or recognizing the new type of knowledge that result from modern data processing (p.321). It was further argued that even when recent or updated laws were effective regarding personal data, they are still not equipped to deal with correlated data, which is persistent today since “(1) group profiles are often inferred from anonymous personal data to which data protection regulation do not apply and (2) group profiles do not necessarily apply to identifiable persons but may, even so, affect the autonomy, privacy, security and equality” of individuals (page.321).

3. Depersonalization.

Depersonalization is a concept in personal data collection that builds on the principle that researchers do not necessarily need the personal identities of data subjects in order to be effective or to achieve objectives because what is only required for legitimate research is statistical access. Ideally, depersonalization renders a data subject completely anonymous. However, this is impossible to achieve in most applications that is why a modified definition was put forward, which states depersonalization as “the modification of personal data so that the information concerning personal or material circumstances can no longer or only with a disproportionate amount of time, expense and labour, be attributed to an identified or identifiable individual” [2] (Fischer-Hubner 2001, page.112). The definition was contained in the groundbreaking Federal German Data Protection Act, which already became synonymous to practical depersonalization. To demonstrate this in real-world application, this paper cites a Lightweight Data Security System developed by Rawassizadeh [9], which provided a working framework that can provide insights how depersonalization actually works (Fig.1). Based on the architecture, the user inputs data into a system that include several stages of pseudonymization before personal information is published or made available to third parties.

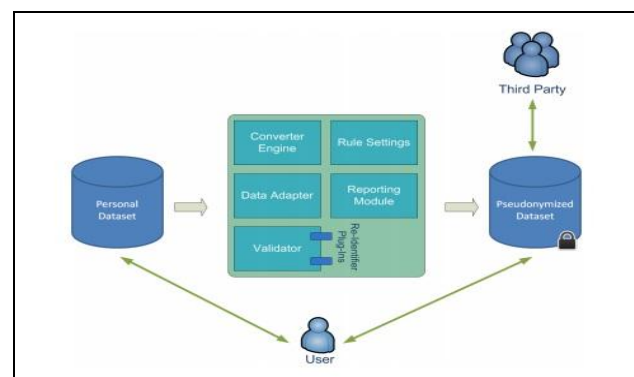


Fig 6. Lightweight Data Security System Conceptual Framework (Rawassizadeh at al .page. 3)

Neubauer and Heurix [6] (2011) further provided specific applications when they proposed a depersonalization system or pseudonymization of medical data to be used in health care institutions. The application is called Pseudonymization of Information for Privacy in e-Health or PIPE, which aims to provide a “traceable anonymity” (page.194). It works using a combination of symmetric and asymmetric cryptographic keys in order to achieve a logical multi-tier hull model composed of three layers

(Fig. 2). The model is applicable in several health care scenarios as shown in Fig. 3.

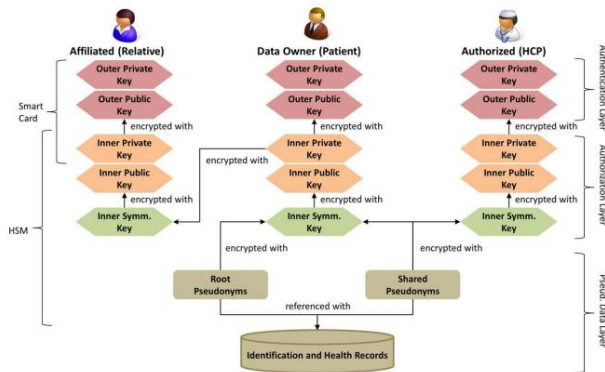


Fig 2. The PIPE Model (Neubauer & Heurix, 195).

Based from the sample frameworks provided, it is clear how the depersonalization system works and how data is stored, maintained and protected. Its adoption will entail the installation of additional application but it will effectively address ethical and legal questions on personal data gathering practices.

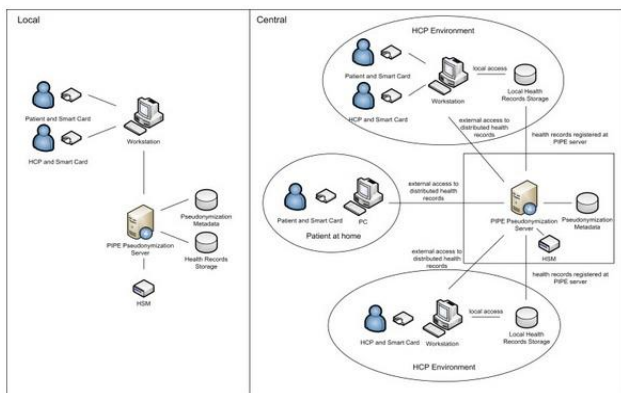


Fig 3. Scenarios for PIPE Application (Neubauer & Heurix, page.195).

4. Challenges and Future Trend

Depersonalization, certainly, is not perfect. The method and its resulting applications still entail risks since the anonymity it could provide is not absolute. This was earlier cited with the modification in the definition of the concept. Actual depersonalization of data, wrote Wagner [11], is weakened by the practice of linking different records to increase their information content and link the outcome to the identity of individuals (page.3). Neumann [8] also argued that the method of depersonalization may lead to temptation to commit misdeeds, diminish human initiative and hinder the principle of accountability, and eroding our sense of ethical behaviour in the process.

There are risks involved such as the methods available to “re-identify” anonymous individuals. But the process, as has been pointed out by the German law cited in this paper, becomes time and resource consuming that undertaking them becomes impractical. There are also mechanisms and applications that can prevent the risk of re-identification. This is particularly important in the area of future work in this field. Like any form of technology, it rapidly evolves, updating capabilities so that there is a potential of more sophisticated and effective models.

5. Conclusions

All in all, depersonalization is an effective and viable solution to personal privacy concerns amidst the tremendous power available to governments and businesses to gather personal data. It is a technical solution to a technical problem. The inefficacies of enacted laws to protect personal privacy serve to highlight this point. Fundamentally, adopting it makes sense because it addresses the problem from the very beginning: the identities of data targets are masked and the statistical information is accessible. It solves privacy issues and satisfies the need for data so that governments and businesses are able to provide products and services that are better and more attuned to our needs.

References

- [1] Bidgoli and Hossein, “The Internet Encyclopedia”. Hoboken, NJ: John Wiley and Sons, 2004.
- [2] Simone Fischer-Hubner and G. Fischer-Hubner, “It-Security and Privacy: Design and Use of Privacy-Enhancing Security Mechanisms ” Issue 1958. Berlin: Springer, 2001. Print..
- [3] Conrad, Eric, Misenar, Seth, and Feldman, Joshua. “CISSP Study Guide”. Waltham, MA: Newnes, 2012. Print..
- [4] Hildebrant, Mirreile and Gutwirth, Serge. “Profiling the European Citizen: Cross-disciplinary Perspectives”. Berlin: Springer, 2008. Print.
- [5] Kluge, Eike. “Professional codes for electronic HC record protection: ethical, legal, economic and structural issues.” *International Journal of Medical Informatics*, 2.1 (2000): 85-96. Print..
- [6] Neubauer, Thomas and Heurix, Johannes. “A methodology for the pseudonymization of medical data”. *International Journal of Medical Informatics*, 80(2011), 190-204. Print..
- [7] Neubauer, Thomas. and Kolb, Mathias. “Technologies for the Pseudonymization of Medical Data: A Legal Evaluation”. 2009 *Fourth International Conference on Systems*, 2009. Print.
- [8] Neumann, Peter. “Consideration on risks in using computing technology.” *ACM SIGSAC Review*, 6.2 (1988), 2-4. Print.
- [9] Rawassizadeh, R., Heurix, J. Khosravipour, S. and Tjoa, A Min. “LiDSec: A Lightweight Pseudonymization Approach for Textual Personal Information.” Vienna: Vienna University of Technology.

- [10]Szeto, M. and Miri, A. "Analysis of the Use of Privacy-Enhancing Technologies to Achieve PIPEDA Compliance in a B2C e-Business Model." *Eighth World Congress on the Management of eBusiness*, 2007. Print.
- [11]Wagner, Gert. "Autonomous Organization of the (International) Scientific Community Would Simplify Data Protection in the Social Sciences and Encourage Reanalysis." *ECONSTOR: Discussion Paper N. 249*, 2001. Print.

Bojken Shehu. He is a pedagogue in Polytechnic University of Tirana, Faculty of Information Technology, in Computer Engineering Department. In 2007 he has finished the Bachelor Thesis in Saint Petersburg State Polytechnical University, Russia. In 2010 he has finished the Master Thesis in Bauman Moscow State Technical University, Russia, with excellent results and now he is a PhD student in Polytechnic University of Tirana, Albania.

Shqiponja Ahmetaj. She is a Master of Science student in Vienna University of Technology, Faculty of Informatics. In 2011 she has finished the Bachelor Thesis in Saint Petersburg State Polytechnical University, Russia, with excellent results.

Marin Aranitasi. He is a pedagogue in Polytechnic University of Tirana. He has finished the Bachelor and Master Thesis in Polytechnic University of Tirana, Albania and now he is a PhD student in Polytechnic University of Tirana, Albania.

Aleksander Xhuvani. He is a pedagogue in Polytechnic University of Tirana, Faculty of Information Technology, in Computer Engineering Department. He has finished the PhD study at Bordeaux in France. At 2004 he is graduated as Prof.Dr.

The Analysis and Design of a Web-Based Social Network: Locate it! Project

Noura AlOmar¹ and Wea'am AlRashed¹

¹Software Engineering Department, King Saud University
Riyadh, Saudi Arabia

Abstract

Locating missing items is a challenging problem that people face frequently. This paper proposes a system that attempts to solve this problem by taking the advantage of the interactions among people in online social media. Locate it! Is an online social community for connecting individuals with lost items with people who found these items. The communication in this Arabic social network is based on the contribution of individuals in the Kingdom of Saudi Arabia. After reporting the lost/found items by the users, they can ask the system to retrieve matches from the database. As a result, this research study discusses the analysis and design of Locate it! System and concludes by summarizing the paper; future improvements are also discussed.

Keywords: Location management, Lost and Found, Localization, community-oriented system, Online Social Network, Web-based system.

1. Introduction

On a daily basis, people face the problem of misplacing their stuff which might be valuable or have a special meaning to them. Also, some items are irreplaceable and despite the fact the other items can be replaced, the users might have a limited time to find a replacement and they don't want to waste their time in searching for them. Another case is when a person loses an item that has very sensitive information and this loss might cause a considerable impact such as misplacing mobile device, identity cards, and passports. Technology has the opportunity to assist people in this problem. In recent years, several systems have been introduced to keep track of lost items and help people in finding them [e.g. 1, 2, 3, 5, 6, 7 and 8]. Approaches have varied from RFID tagging to community based systems which allow the work and cooperation between the members of the virtual community to facilitate the process of locating the missing objects through exploiting the capabilities of the Internet.

Moreover, anecdotal evidence from observations in public areas shows how individuals often seek the assistance of others by posting announcements such as the notes depicted in figure 1 for missing items. In this example, the signs were posted by the owners of lost items who offered incentives to encourage and elicit responses and assistance from the community. On the other hand, people always find lost things everywhere and they want to deliver them to their owners who they don't know.

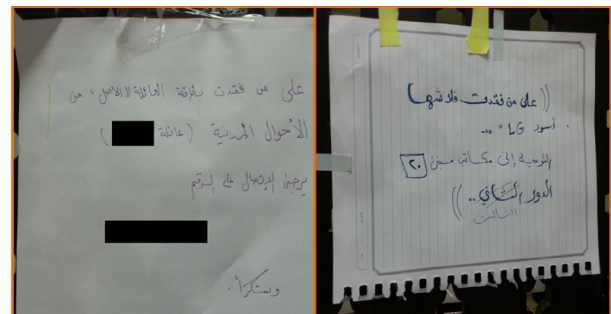


Fig. 1 Lost items announcements

From all the above, it is apparent that there is a lack of a system that support Arabic language and can do the job of locating missing items and return them back to their owners within the context of Saudi Arabia. We also have noticed the high level of participation and interactivity between the members of our society in social networks [4] like Twitter, Facebook, Instagram and how the people are willing to help each other through these networks. In addition, the growth rate of the social networks usage in Saudi Arabia is high. To take Twitter as an example and as reported by The Social Clinic consultancy website [9] "Arabic is the fastest growing language on Twitter" and "Saudi Arabia grew more than 3,000% from 2011 to 2012, and currently accounts to 50,000,000 tweets per month".

There exists inadequate support of Arabic-language based online communities dedicated to the lost and found dilemma. Therefore, designing a social network for the lost and found items in Saudi Arabia where people can interact with each other easily can leverage the usage and participation of the local community in social networks. This project will start to cover the lost/ found items in Riyadh and our future vision is to let it includes the entire Kingdom of Saudi Arabia after examining the behavior of the users and knowing how to make the appropriate improvements. The concept aims to automate the process of connecting the person who has lost something with the person who has found something in a cost-effective, enjoyable and simple way to help them in saving their time and effort.

This paper is structured as follows: Section II looks at the literature and describes the work related to locating the missing items and connecting the owners of lost items with individuals who have found these items. Section III discusses the analysis of the distributed questionnaire. The following section provides the basis that will be followed in the matching algorithm. Section V illustrates the diagrams that are used to model the system and Section VI gives two prototypes of the system's interface. We conclude in section VIII with a summary of what has been introduced in this research study along with the future work of the system.

2. Literature Review

In recent years, advances in technology in general, and social computing in particular, have led to proliferation in research examining how these networks can facilitate better communication amongst individuals [e.g.10]. In the context of the lost and found for personal possessions, there are few research studies which have been made for the purpose of locating missing objects, however; most of them were concerned with detecting the exact location of the missing object through using technologies such as Radio Frequency Identification (RFID) readers, ultrasonic position detection systems and sensors [5,6 and 7]. There are three examples of such research studies, first one is a support system for finding lost objects using spotlight [5] which were developed to locate the objects indoors using light notification technique. Another example was the system named: Where's My Stuff? [6] and was designed for helping visually impaired individuals to find their missing objects within a range of 30 meters through using Bluetooth-enabled tags which can be activated via an application installed on devices that have Bluetooth technology. The third one was mainly relying on a concept

common to our proposed idea in that it adopts a community-based reporting system to trace the lost/found items. On the other hand, it relies on using RFID readers which are embedded in mobile phones to scan objects' RFID tags and send the gathered data to the server-side software to find the objects' owners [8]. Furthermore, [7] provides a full analysis of the type of lost objects, places where they have found, persons who have lost them and the search strategies which were used based on distributed surveys. In addition, through our search for similar systems, three websites have been found which are developed to provide services similar to our proposed idea. These websites are: Lost my stuff [1], The Internet Lost and Found [2] and Mafgod [3][Figure 2]. As for the language of those websites, the first two websites are using the English language. However, the third one is an Arabic website and by comparing them to our system, we are planning to use the Arabic language as the main language and we are going to provide the option to use the English language if necessary. Moreover, Mafgod is dedicated to serve people who have lost or have found mobile phones only. In contrast, the Locate it! System is not limited in serving a small segment of the society; almost all kinds of items can be reported as lost or found without any limitations. Furthermore, Lost my stuff [1] website is designed to rely on volunteers who are registered on the website and have specified their locations. Thus, whenever an item is reported as lost, the system forwards the request to a volunteer who exists in a location near from where the item was lost to provide help in finding it. As a result, it doesn't use an algorithm to match items with each other and generate recommendations. Also, it doesn't have the functionality where anyone can report a found item. On the other hand, we are planning to implement a recommender system which is based on the comparison of the lost items versus the found items which are stored in the database and generate recommendations to assist the user in locating his/her items in a short time.



Fig. 2 The three websites

In addition, we can consider The Internet Lost and Found [3] website to be the most similar system to our system. This website and our system have a common functionality which is allowing the registration of places that are sometimes relevant to specific events. For example, registering conferences, seminars, workshops, restaurants, malls ... etc. Thus, you can think of our system as a big box which acts as a container of many small boxes.

As best of our knowledge and from the comprehensive research that has been done so far, there is a lack of an Arabic interactive system which serves people in recovering their belongings and the purpose of our system is to bridge this gap. Furthermore, to increase the user satisfaction and to simplify the process of locating the lost items, we will make our system available as a website and as a mobile application. We chose to implement two versions of the site, an online web based version and another mobile application because many of our target users are users of smart phones and it is more efficient in reporting an item as either lost or found to make the process of helping people in finding their items faster.

3. Survey

As a part of this study, a survey has been developed and it was intended to explore the requirements of our systems and to gain a sufficient understanding of the intended users' desires. This survey was created and distributed online. Our sample of respondents consists of 75 individuals recruited from a convenience sample; all of them were Saudi Arabian nationality. The survey was comprised of questions to explore patterns of user behavior and their

experiences in losing items and trying to relocate them. It also examined the demand of creating a social network for the purpose of letting people share their missing and found things.

The analysis of the survey shows that 89% of the respondents specified that they need an automated system to help them in locating their missing stuff and 78% stated that they have lost something in the past and they didn't find it yet. These results confirm our idea that there is a gap which needs to be filled. In addition, 16% of our sample participants said that they have made announcements about their lost stuff in the social networks. Although this ratio is quite small, it is acceptable due to privacy issues, since some people don't want to disclose such information to be available to the public in social networks because there is a possibility to be stolen by other persons whom might search for it and find it faster [Table 1]. Furthermore, the questionnaire has listed 19 item types and has asked the respondents to indicate the item types that have been missed before. The goal of this question is to pinpoint the item types that need to be existed in add lost/found item form (Section VI). The answers of this question show that 50% of the sample have lost their mobile phones before, 45% have lost their flash memories, 39% have lost their keys and a ring, 33% have lost a bracelet, 24% have lost a necklace, 23% have lost a wallet and a watch, 17% have lost a book, 15% have lost a notebook, 12% have lost a bag, 11% have lost their ID's and a hard disk, 8% have lost their iPods, 7% have lost their laptops and iPads, and 4% have lost their passports [Table 2]. Moreover, it is important to note that our system will have a service that allows the admin to add other item types depending on the users' needs.

Table 1: Survey's Yes/No questions

Question	Yes	No
Do you need an automated system to help you find your lost items?	67 respondents	8 respondents
Did you lose an item in and haven't find it yet?	59 respondents	16 respondents
Did you use social networks to announce for a lost item?	12 respondents	63 respondents

Table 2: Item types

<i>Item type</i>	<i># of respondents who have encountered this loss</i>
Passport	3
ID card	8
Mobile phone	37
Keys	29
Bag	9
Laptop	5
Wallet	17
Ring	30
Bracelet	25
Necklace	18
iPad	5
iPod	6
Tablet device	1
Watch	17
Game	8
Book	13
Notebook	11
Flash memory	34
Hard Disk	8

On the qualitative section of the survey, there is an area that allows respondents to write additional comments which might give the target users the chance to specify additional functionalities to be added to the system. Some of them want the system to employ multiple languages, other respondents indicated the need to provide a delivery service to deliver the located items to their homes, and others want the system to serve people with reading disabilities by pronouncing the words written in the user interface. Finally, some wrote that the fast response from the system is the most important service that leads to a successful project.

4. Matching Algorithm

As described in the previous sections, our project is mainly dealing with three types of stakeholders:

- 1- The person who has lost something.
- 2- The person who has found something.
- 3- A person from a venue side, who is responsible for keeping track of the lost and found items for a specific place such as restaurants and airports.

So it is clear that there is a matching involved in the process of pairing the lost item with the found one. As a result, this section describes how the Locate it! system will perform the matching process and the basis which we are going to follow in writing the source code. To make the

result of the matching process accurate and to increase the possibility of reaching the correct results, there are several characteristics of the lost/found items which were taken into consideration when implementing this project, these factors are:

- 1- The brand of the lost/found item.
- 2- The type of the lost/found item.
- 3- The place [City and street] where the item was lost or found.
- 4- The color of the lost/found item.
- 5- A photo of the lost/found item.
- 6- The material of the item, i.e. metal, leather,..
- 7- Additional specifications if there are marks which uniquely identify the lost/found item (if applicable).

After asking the user to provide the system with the required information of the lost/found item and storing them in the database of the system, the system is going to search about an item with similar characteristics by comparing the factors of the added item with the lost/found items which are stored in the database depending on its kind. For example, if the added item is lost, the system is going to search about a match in the found items database and vice versa. In addition, if the system finds one or several similar matches of a lost item, the system is going to present their photos to the person who found it. Therefore, if the founder confirms one of these pictures and finds that it is the same as what he has, the system will offer two choices for him. The first choice is to give him the contact information of the person who owns this object and they will contact each other either by phone or email to deliver the item. The second choice is to make our system as a broker between the two parties, and the system administrators will take the responsibility of delivering this item to his owner and we may deal with one or several shipping companies for this purpose.

5. System Models

As a part of the analysis and design phases of the system, it is essential to draw graphical representations to demonstrate the structure of the system. In this paper, DFD Context Diagram and Use Case Diagram have been chosen to represent the system. The following sub sections provide a high level modeling of our system and thus they represent the system as a set of functions, the input/output for each one and the actors involved.

5.1 DFD Context Diagram

The structural approach is followed to model Locate it! System because this system is better described by functions rather than classes. Thus, figure 3 conceptualizes the system by showing the entire system as a single process. Moreover, it represents the information exchange between the actors from the external environment and the system.

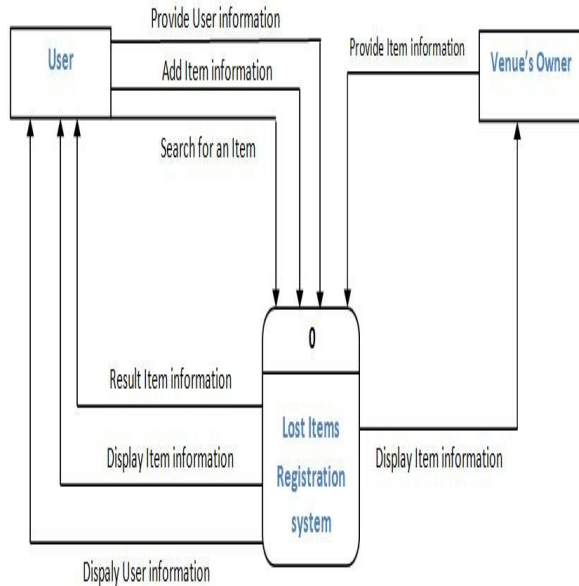


Fig. 3 DFD Context Diagram

5.2 Use Case Diagram

In order to clarify and organize the system's requirements by describing "who" can do "what" with the system, the Use Case Diagram is provided [figure 4]. This diagram shows the main system functions that can be done by each actor of the system. These actors are the Venue Owner, the User which can be the person who has lost or found an item and the last actor is the Administrator which is responsible for managing the system.

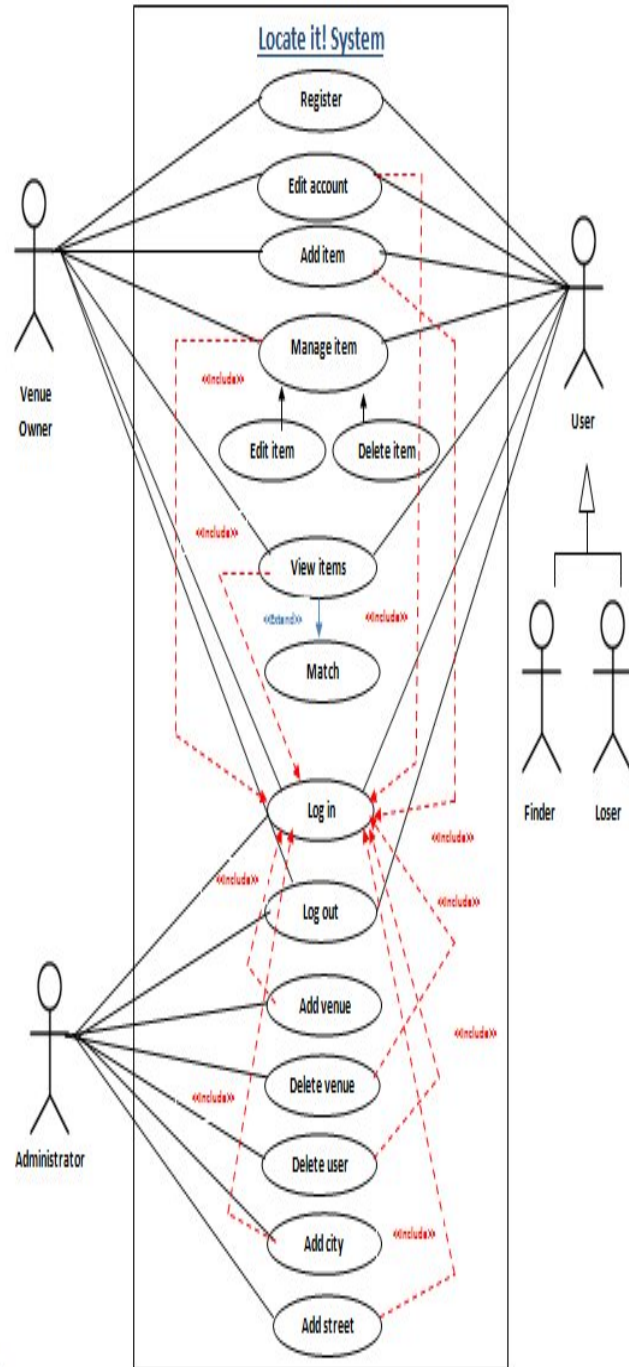


Fig. 4 Use Case Diagram

6. Prototype

A prototype for the Locate it! system has been developed. The prototype is comprised of two forms, one is the registration process form and the other one is for adding a lost/ found item.

6.1 Registration Form

In the below registration form the user has been asked to provide his personal information so the people won't have problems when trying to communicate with him.

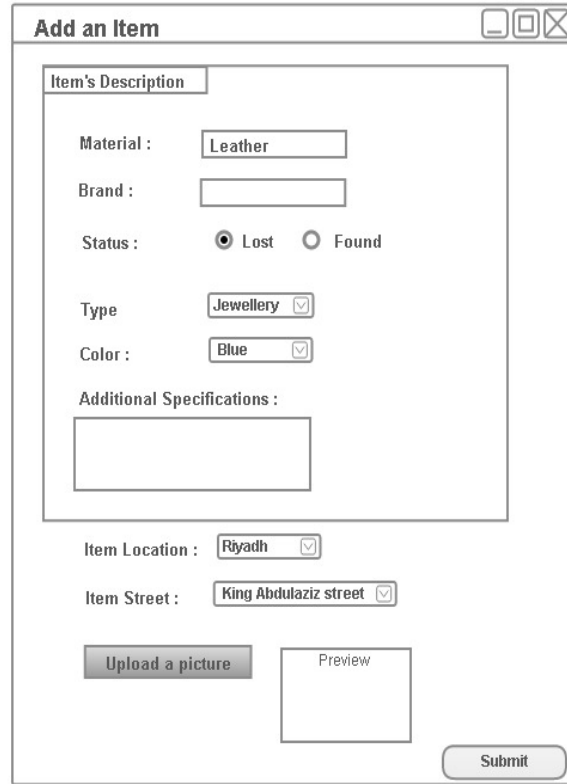


The registration form is titled "Registration" and includes a welcome message: "Welcom To Missing stuff web site ... Please provide the following information so we can be able to communicate with you." The form contains the following fields: "First Name", "Last Name", "Email Address", "Mobile Number" (with an example "055xxxxxxx"), and "City" (with a dropdown menu showing "Riyadh"). A "Submit" button is located at the bottom right.

Fig. 5 Registration form

6.2 Report Item as Lost/Found Form

In "add an item" form, the user is asked to provide different item's information and indicate the status of it whether it was lost or found, this is important to classify the item's situation when the user request for lost/found items. And thus, display it in an organized manner. In addition, the item's location and street were required to be filled by the user in order to organize the items based on their location. This will simplify the process when searching for a lost/found item which is located in specific place. Furthermore, the user has the ability to enter any additional information in the "additional specifications" text area and also can provide an image for the item by pressing the "upload a picture" button.



The "Add an Item" form includes a title bar and a "Submit" button. It features a large "Item's Description" box containing: "Material" (text input: "Leather"), "Brand" (text input), "Status" (radio buttons: "Lost" selected, "Found"), "Type" (dropdown: "Jewellery"), and "Color" (dropdown: "Blue"). Below this is an "Additional Specifications" text area. At the bottom, there are "Item Location" (dropdown: "Riyadh") and "Item Street" (dropdown: "King Abdulaziz street") fields. A "Preview" box and an "Upload a picture" button are also present.

Fig. 6 Report an item as lost/ found form

7. Conclusion

By relying on the contribution of people in online social networks, the idea of this project has been proposed to help people in finding their missing objects by creating an Arabic online social community where people can easily report their lost and found items. Moreover, the ability of creating small virtual boxes for venues such as shops and seminars is supported by this system. This project will start to cover the areas in Riyadh and will expand to cover the Kingdom of Saudi Arabia. This paper introduces the analysis and design of this interactive environment. We have looked at the literature to find how the problem of locating missing items has been solved previously and compare our system's functionality with the existing systems. A survey has been collected from seventy five participants to explore their requirements and to know their experiences in missing their objects. In addition, the use case diagram and the context diagram have been mentioned in this paper to model the system and represent its functions and actors. Furthermore, two prototypes of the system have been provided which are the registration form and the report an item as lost or found form. The main purpose of this project is to take the advantage from the

dynamic data in social networking to create a community of people helping each other and collaborating for locating the objects in a secure and reliable way.

References

- [1] Lost my stuff, Place: [Online]. Available: <http://www.lostmystuff.net/index.php>, Year 2010.
- [2] The Internet Lost and Found, Place: [Online]. Available: <http://www.lostandfound.com/>, Year 2013.
- [3] R. Mohammed , Mafgod, Place: [Online]. Available: <http://www.mafgod.com/>, Year 2006.
- [4] D. M. Boyd and N. Ellison, "Social network sites: definition, history, and scholarship", *Engineering Management Review, IEEE* , Vol.38, No.3, Year 2010, pp. 16-31.
- [5] N. Toyohisa, K. Hideaki, and K. Susumu , "A support system for finding lost objects using spotlight", In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services (MobileHCI '05)*, Year 2005, pp. 321-322.
- [6] K. Julie , P. Shwetak , T. Arwa , G. Brian , W. Jennifer , and A. Gregory , "Where's my stuff?: design and evaluation of a mobile system for locating lost items for the visually impaired ", In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*, Year 2006 , pp. 103-110.
- [7] Rodney E. Peters, Richard Pak, Gregory D. Abowd, Arthur D. Fisk and Wendy A. Rogers, "Finding Lost Objects: Informing the Design of Ubiquitous Computing for the Home".
- [8] G. Dominique , B. Oliver , and M. Florian , " Supporting a mobile lost and found community ", In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services (MobileHCI '08)*, Year 2008 , pp. 407-410.
- [9] The State of Social Media in Saudi Arabia, Place: [Online]. Available: <http://www.thesocialclinic.com/the-state-of-social-media-in-saudi-arabia-2012-2/> , Year 2012.
- [10] M. Marti and V. George , " I seek you: searching and matching individuals in social networks ", In *Proceedings of the eleventh international workshop on Web information and data management (WIDM '09)*, Year 2009 , pp. 67-75.

Noura AlOmar has a bachelor degree from the College of Computer and Information Sciences in the field of: Information Technology, King Saud University (KSU), in 2011. Noura's current job is a teaching assistant in the department of Software Engineering, College of Computer and Information Sciences at King Saud University, Riyadh, Saudi Arabia. Ms. AlOmar has few international academic publications which include conference papers and journal papers. In addition, she has published few articles in local newspapers. Her research interests include recommendation systems, community-oriented systems, Web technologies and Bio-Informatics. Furthermore, her research areas in human-computer interaction cover the user-experience testing methods and techniques, affective computing and exploring users' behaviors in online social networks.

Wea'am Al-Rashed has a Bachelor degree in information technology from college of computer and information science at King Saud University at the year 2011 in Riyadh , Saudi Arabia. Wea'am's current job is a Teaching Assistant in the Software Engineering department at King Saud University. Wea'am had published three papers in two different conferences at 2010 and two other papers in two journals at 2013 .Wea'am's research focused on Web development, Research engines, Software engineering and human computer interaction (HCI).

An Eccentric Scheme for Oblivious Communication

Khan Farhan Rafat and Muhammad Sher

Department of Computer Science, International Islamic University
Islamabad, 44000, Pakistan

Abstract

Trust is the foremost requirement in any form of communication. Building trust through physical contact, gesture etc. is easy but hard to establish in electronic data communication as one can't be sure of the presence of the intended recipient at the other end. This necessitated the need to devise and suggest covert schemes for oblivious communication that only the intended recipient may unveil. Steganography is one such technique where information to be sent is seamlessly superimposed on the carrier such that it easily crosses the information confronting barriers without detection. This paper endeavors at evolving a data hiding scheme envisaged by Shannon and in adherence to Kerchhoff's principle that hides secret information inside statistically random English alphabets appearing as cryptogram. Effect of bit embedding is computed through contrasting probability distribution of cover text and stego object.

Keywords: *Information Theoretic Security, Cryptogram, Trust, Oblivious Communication, Steganography.*

1. Introduction

400 B.C. saw Greeks articulating Steganography for covert communication. Herodotus while accounting on the quarrel of 5th Century B.C. between Greece and Persia elaborated in [1] on how Demaratus; exiled from Greek; informed his natives in Persia of Xerxes plan of invading Greece by engraving secret message on wooden tablet and coating it with wax where recipient melted wax to read the carved message.

In another incident, as narrated by Herodotus; another Greek named Histaiacus; called for an insurgency against Persian King by shaving and tattooing a secret message of revolt on the head of one of his slaves. Ample time was then given for his hair to grow so as to hide / cover the tattooed message after which he was sent to the enemy territory. The intended resident in Persia read the hidden message after shaving the head of the messenger.

Ancient Chinese are known to write their message on a piece of silk which they crunched into a little ball. The ball is then covered with wax and gets swallowed by the messenger [2].

Æneas, the author of the book titled "On the Defense of Fortified Places" developed a Steganographic System by blurring holes into wood that corresponds to Greek

alphabets and then threading the yarn through these holes in an order which when detached spelt out letters of hidden message in its correct sequence [3]:

"padielaporsymesarponomeuaspeludynmalpreaxo"

"Ave Maria" is yet another smart method for encoding secret message where list of distinct words for each unique alphabet are grouped in a series of Tables. For every alphabet constituting the message, word corresponding to that letter from the set of Tables gets substituted and appeared as innocent rhyme whenever the Tables are used in sequence.

Cardano introduced the "Grille System" [4] where every recipient was given a piece of paper with a number of punched holes in it which when placed over an innocuous Stego Object (Text Message) reveals letters of hidden secret message.

In his book "Mercury" often referred to as "The Secret and Swift Messenger", author Bishop John Wikims described ways such as use of alum, ammonia salt, onion juice besides "distilled juice of Glowworms" for hidden writing that glows when exposed to light [5].

In an attempt to dethrone protestant queen Elizabeth, Mary a Catholic, made use of information hiding techniques to secretly communicate with Catholic Noblemen of England [6].

George Washington (General and first United States President) also made use of Steganography and Cryptography for cover communication and ordered for the use of invisible inks on regular messages in an attempt to divert suspicion [7].

Franco-Prussian war (1870-71) saw Pigeons brought in France, 1000 in numbers, being used as message carriers for delivering innumerable Military and more than 95,000 secretive dialogues from and into France [8].

Tobin et al. in [9] elaborated on the use of Quilts, hung outside to dry, with special symbols / patterns sewed into its body thereby innocuously pointing towards directions to help prisoners escape.

World War – I (WW-I) saw improvement in Cardano's Grill that gets rotated 90 degrees each time a series of letters get encoded when Germans introduced different type of Grills based on varying message lengths. This

technique, however, was cracked within four months since its inception. In another incident a lady alleged as German spy was found in possession of blank sheet of paper which when inspected exposed a hidden message written with invisible ink [10].

Radio communication was also used for the purposes of message transmission, for the first time, also during WW-I. Cable censor and Cigar movements conforming to ships movements are also some popular examples of the same retro [11].

World War II also witnessed extensive usage of Steganography as explicated through text messages written in quotes using technique called concealment cipher, extracted from [12] as under:

“Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on by products, ejecting suets and vegetable oils.”

The stealth message “Pershing sails from NY June 1.” can be extracted from afore mentioned quoted string by taking second letter of each word.

Captivated crew members of U.S. Naval spy ship “Pueblo” used body gestures to secretly communicate the message “Snow Job” when the ship got captured by North Koreans on January 23, 1968 [13].

During Vietnam War (1954) Commander Jeremiah Deuton of United States Navy secretly communicated the plight “T-O-R-T-U-R-E” through his eyes using Morse code [14].

Former British Prime Minister Margret Thatcher was so aggravated with information degeneracies that she ordered for Word Processors to secretly hide identity of its user in word spaces to identify the disloyal amongst her Ministers and Staff [15].

Aligned with context of this paper there exist a web-based text generating Steganographic tool called “spammimic” known for using “spam” grammar and a mimicry algorithm named after Wayner to generate text cover [1].

The most recent case of hidden communication is the famous verdict by Judge in the “Devince code” case where first few lines of his verdict contains his hidden nickname as “Smithy Code” i.e., a statement given by Smith [16].

Prior to proceeding further it may not be out of place to mention here some personalities who presented revolutionary ideas often emphasized in context of cryptography but surely are rudimentary to ensure security of any cryptographic or Steganographic system:

- a. Born in 1535 Giovanni Porta gave the concept of substitution and transposition that earned him a name in the history of Cryptography [17].

- b. Auto Key concept given by Blaise de Vigenere that got re-invented in the nineteenth century [18].
- c. According to Kerckhoff security of a System must be a feature of its Key when the algorithmic details are made public [19].

Rest of the paper is premeditated as follows: Section 2 builds the foundation for our proposed solution which is explained in detail in Section 3. Section 4 illustrates test results. Advantages and recommendations are given in Section 5 and 6 respectively. Section 7 concludes the discussion.

2. Building Blocks

Text Steganography uses “text” as cover for secret communication. Operations such as text formatting, text insertion, text substitution, generating text through context-free grammars etc. are in use for text Steganography. Further “text format” stands alone when it comes to saving extra information (meta data) associated with specific file format i.e., text files are saved, retrieved and viewed as these appear before human eye, and it is this peculiar trait that makes it difficult to devise ways for hiding information inside text cover.

Images (due to high data capacity) by far are the most preferred choice when it comes to cover selection for the purpose of information hiding. However, as in [20], [21] and [22] major limitations of Image Steganography include:

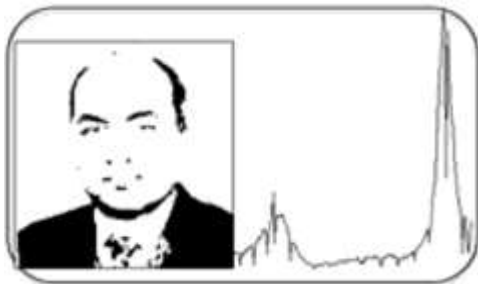
- a. Data/Information Loss when compressed using lossy compression.
- b. Choice of cover image.
- c. Original cover or its related information is required at receiver's end for information retrieval.

The popular method for image Steganography is inserting secret message bits at least significant bit positions in the cover image, called as LSB (Least Significant Bit) Steganography. Figure 1 illustrates hard to distinguish (from perceptibility point of view) cover and Stego images along with their graphical representation respectively.

2.1 Difficulty in using LSB scheme for ASCII Text Steganography

To highlight the difficulty associated with hiding data inside text cover, we chose the cover text (extracted from <http://ijcsi.org/about.php>) shown in Figure 2 and substituted “0” and “1” at LSB position of each of its words the outcome of which are illustrated in Figure 3 and 4 respectively.

It is evident from above illustrations that even a single bit change at LSB position in cover text raises perceptibility concern and hence may not be used for information hiding.



Cover (image) with histogram



Stego Object (image) with histogram

Fig. 1 Image Steganography – Cover Image and Stego Object

2.2 Frequency Count Reveals Data Hiding Scheme

Samuel Morse (1791-1872) contribution of encoding English alphabets based on their frequency of occurrence [23] in Standard English text paved the way for American Standard Code for Information Interchange referred to as ASCII character set and pronounced as /'æski/ASS-kee [24] for digital information exchange / electronic communication. The concept of use of frequency of occurrence, however, has been browbeaten by cryptanalysts in solving cryptogram i.e., *counting frequency of occurrence of English alphabets helps in revealing the type of encryption technique being used as either substitution or transposition* [25]. Case insensitive analysis of English alphabets for cover text gives the frequency distribution shown in Table 1.

Frequency count of English alphabets shown in Table 1 seems in agreement (to a great extent) with that given in [12] and illustrated in Figure 5, from where cryptanalyst may easily infer that the text under examination is the outcome of substitution scheme.

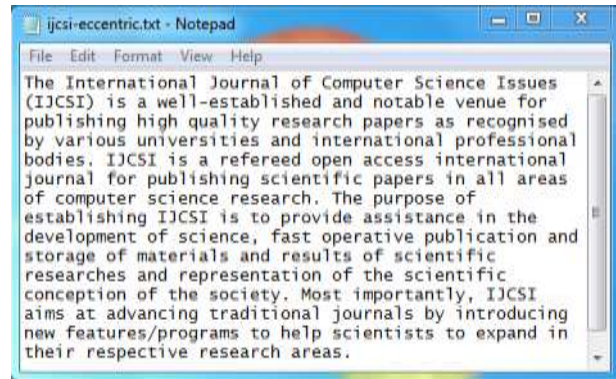


Fig. 2 Chosen Cover Text

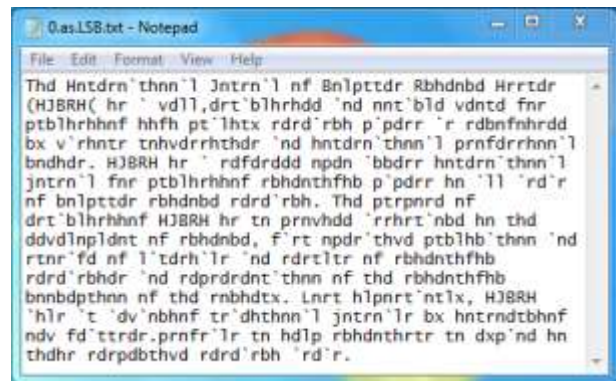


Fig. 3 Cover Text with bit-0 at LSB position of each word

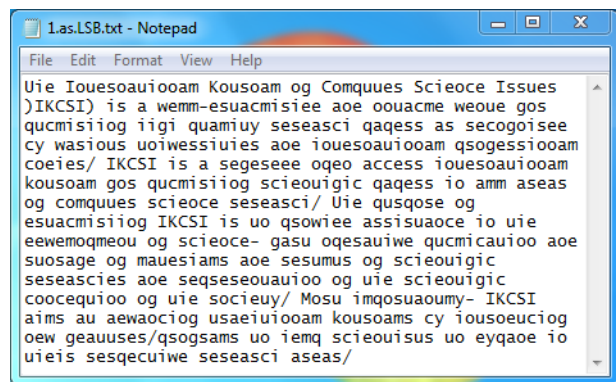


Fig. 4 Cover Text with bit-1 at LSB position of each word

Assuming the under discussion text be a result of mono-alphabetic substitution, even then the frequency of substituted letters would correlate with that as in Figure 5, thereby enabling the cryptologist to infer that before him is a case of substitution cipher.

2.3 How to Impede Prediction?

Aforesaid revealing attribute of English alphabets resulted in use of cryptographically secure randomly generated numbers through hard to predict processes like radioactive decay, thermal, acoustic sources [26] etc. Randomness

adds uncertainty that elevate Wendy’s effort towards determining the underlying information hiding scheme – a concept also supported by Information Theory.

Table 1: Frequency Count for English Alphabets in Cover Text

Alphabets	Frequency	Alphabets	Frequency
A	54	N	51
B	9	O	43
C	34	P	23
D	15	Q	1
E	78	R	44
F	17	S	61
G	9	T	48
H	17	U	17
I	69	V	8
J	7	W	2
K	0	X	1
L	24	Y	5
M	8	Z	0

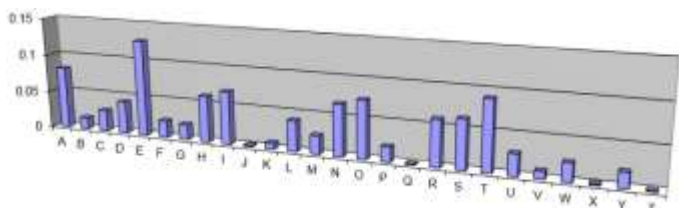


Fig. 5 Frequency Count For English Alphabets Based On Their Occurrence In Standard Text

2.4 Cryptogram

Electronically generated output of encryption process called cryptogram can take any form but are best expressed in terms of alphabets, numbers or blend of two and takes care of the secrecy aspect of communication security amongst confidentiality, integrity, and availability (C.I.A.). Figure 6 is an illustration as on p - 99 in [12]. For more recent example [27] is referred.

The desirous attribute of Cryptogram is to exhibit unpredictability/randomness [28], and if possessed is then ascertained to pass statistical tests evolved to verify on the randomness of generated data [29].

This section suggests a cover generation scheme envisioned from Shannon’s deliberation on types of secrecy systems [30] where he conferred that furtive

systems may exercise methods such as “..., message in fake cryptogram, ...”, and discussed as follows:

3. Proposed Atypical Steganographic Scheme

Perceptibility issue of Stego objects (Figure 3 and 4 refers) as a result of LSB substitution of words in Cover text (Figure 2 refers) ensued our analysis of various combinations of 7-bit binary strings for ASCII codes (with zero prefix) so as to arrive at the best possible scheme where LSB substitution does not lead to non-alphabetic code thereby avoiding perceptibility concerns together with least programming constraints in its subsequent implementation. The study, however, remained confined only to upper case English alphabets, where Table 2 illustrates our findings.

GJXXN	GGOTZ	NUCOT	WMOHY	JTKTA	MTXOB	YNFGO
GINUG	JFNZV	QHYNG	NEAJF	HYOTW	GOTHY	NAFZN
FTUIN	ZANFG	NLNFU	TXNXU	FNEJC	INHYA	ZGAEU
TUCQG	OGOTH	JOHOA	TCJJK	HYNUV	OCOHO	UHCNU
GHHAF	NUZHY	NCUTW	JUWNA	EHYNA	FOWOT	UCHNP
HOGLN	FQZNG	OFUVC	JNZJHT	AHNGG	NTHOU	CGJXY
OGHTH	ABNTO	TWGNT	HNTXN	AEBUF	KNFYO	HHGIU
TJUICE	AFHYN	GACJH	OATAE	IOCOH	UFOXO	BYNFG
GJXXN	GGOTZ	NUCOT	WMOHY	JTKTA	MTXOB	YNFGO
GINUG	JFNZV	QHYNG	NEAJF	HYOTW	GOTHY	NAFZN
FTUIN	ZANFG	NLNFU	TXNXU	FNEJC	INHYA	ZGAEU
TUCQG	OGOTH	JOHOA	TCJJK	HYNUV	OCOHO	UHCNU
GHHAF	NUZHY	NCUTW	JUWNA	EHYNA	FOWOT	UCHNP
HOGLN	FQZNG	OFUVC	JNZJHT	AHNGG	NTHOU	CGJXY
OGHTH	ABNTO	TWGNT	HNTXN	AEBUF	KNFYO	HHGIU
TJUICE	AFHYN	GACJH	OATAE	IOCOH	UFOXO	BYNFG

Fig. 6 Standard Cryptogram Format

Additionally we wanted our proposed scheme to observe the following:

- Kerchoff’s Principle.
- Does not require original cover text at receiver’s end for hidden bits extraction.
- Provisioning for scanning of Stego object.
- Ease in erroneous code detection and its subsequent correction.
- Transmission of Cryptogram in printed form.

Table 2: Considerations before Contriving LSB Based Cover Generation Steganographic Scheme

ASCII code	MSB		Bit String				LSB		Observation	Findings
	7	6	5	4	3	2	1			
65	1	0	0	0	0	0	0	0	64 < 65	Changing LSB to "O" results in "@" < "A"
	1	0	0	0	0	0	0	1	64 = 65	No dispute for bit "I" of secret message
.	1	1	0	0	0	0	0	0	96 > 90	String having its 6 th bit ON are not suitable
	1	0	1	0	0	0	0	0	.	No dispute
.	1	0	1	1	0	0	0	0	.	
	1	0	1	1	1	0	0	0	92 > 90	Perceptibility concern
.	1	0	1	1	0	0	0	0	88 < 90	No dispute
	1	0	1	1	0	1	0	0	90 = 90	
90	1	0	1	1	1	0	1	0	93 > 90	> "Z"
	1	0	1	1	1	1	0	0	94 > 90	2 bit toggling involves more programming constraints with less range of alphabets for use in Cover text
.	1	0	1	1	0	1	1	0	91 > 90	

Bit embedding and extraction processes are sub sequentially discussed as under.

3.1 Bit Embedding

We generated random numbers using a PRNG the discussion of which is beyond the scope of this paper, in range 0 ~ 65535. The same were reduced to modulo 26 and the value 65 (ASCII code for alphabet "A") was added to that yielded an English alphabet in range "A" to "Z" using equation:

$$r \leftarrow \text{random MOD } 26 + 65 \quad (1)$$

Secret message was translated into equivalent binary bits. A 256-bit stego key was fed as input to 256-bit SHA-2 [31] HASH algorithm, the outcome of which was translated into equivalent binary bits and number of ON binary bits (i.e. 1's) counted. *Output from HASH algorithm was fed-back as next input (in place of stego key) as long as the number of ON bits remained less than the number of message bits.* Next, random alphabets greater than or equal to the number of output bits from HASH algorithm (in multiple of 5) were generated and binary bit 1 or 0 from that HASH output was placed sequentially over each alphabet thus generated. Finally the random alphabets were traversed taking each binary message bit, searching for alphabets having an ON HASH bit above it and by skipping the alphabets A and Z. Upon finding the desired random alphabet, secret message bit got replaced with its LSB and substituted with alphabet corresponding to new binary equivalent. *Whenever number of random alphabets fell short of secret message bits new alphabets got generated using aforesaid process.* The process of bit embedding is illustrated in Figure 8.

Algorithm 1.0 – Bit Embedding Algorithm

Function *Eentric_embedding*

Inputs:

x:: String of Random Alphabets,
y:: String of Bits,
Stego_key [256]:: String of 256-bits;

Output:

z :: Output bits; /* Stego Object */

Begin

```
1  q ← Length(x)
2  lm ← Length(y)
3  m ← 0
4  n ← 0
5  for w ← 1 to q
```

Begin

```
6  if Stego_Key(m) = 1 then
7  if Mid(x,w,1) <> "A" and Mid(x,w,1) <> "Z" then
    Mid(z, w, 1) ← Character (ASC(left_7(Mid(x,w,1)) ||
    (Mid(y, n, 1)));
8  n ← n + 1
9  if(n ≥ lm) then exit For Loop
10 m ← (m+1) MOD 256
```

End

End

The complexity of the algorithm 1.0 for worst case is worked out as $O(q)$; while for best case it is $O(lm)$.

3.2 Bit Extraction

Stego Key was fed into 256-bit HASH algorithm as input and the output got translated into bits. The output was also used as next input to the HASH algorithm for subsequent iterations– the process remained unremitting till the output equals alphabets in Stego Object. The ones and zeros thus obtained from Stego Key's HASH were placed over every Stego Object's alphabet.

The LSB of alphabets immediately beneath binary bit 1 of HASH bits less alphabets A and Z were extracted, and segregated into chunks of eight bits. The 8-bit binary strings were translated into equivalent ASCII codes (of English alphabets) that comprised hidden message. Figure 9 illustrates bit extraction process.

Table 3 succinctly exemplifies on bit embedding and extraction process.

Table 3: Eccentric Steganographic Scheme in Nut Shell

Cover Text	ASCII Code	Binary Equivalent	Stego Key Bits	Message Bits	Change in Binary Bits	Changed ASCII Code	Stego Object	Binary Equivalent	Stego Key Bits	Extracted Bits	Hidden Message
H	72	1001000	0		None		H	1001000	0		93= "]"
E	69	1000101	1	1	None		E	1000101	1	1	
L	76	1001100	1	0	None		L	1001100	1	0	
L	76	1001100	1	1	1001101	77	M	1001101	1	1	
O	79	1001111	1	1	None		O	1001111	1	1	
X	88	1011000	1	1	1011001	89	Y	1011001	1	1	
W	87	1010111	1	0	1010110	86	V	1010110	1	0	
O	79	1001111	0		None		O	1001111	0		
R	82	1010010	1	1	1010011	83	S	1010011	1	1	
L	76	1001100	1	1	1001101	77	M	1001101	1	1	
D	68	1000100	1	0	None		D	1000100	1	0	Bits = 10
X	88	1011000	0		None		X	1011000	0		
...

4. Test Results

Test results speaks high on our proposed logic where these were quantified using Hamming [32], Levenshtein (Edit) [33], and Jaro-Winkler [34] distance given in Table 4, mean, variance and standard deviation shown in Table 5 and graphical illustration of Figure 7 using MiniTab 16 [35] by contrasting cover text vs. stego object as follows:

Table 4: Computed Hamming, Levenshtein and Jaro-Winkler Distance

Distance	Computed
Hamming	52
Levenshtein (Edit)	52
Jaro-Winkler	0.861702127659574

Table 5: Computed Mean, Variance and Standard Deviation Before and After Bit-Embedding

Computing	Before Embedding	After Embedding
Mean	0.0773	0.0773
Variance	0.0820	0.0823
Standard Deviation	0.0091	0.0091

5. Advantages

Following justifies to opt for letter based cryptogram solution which, however, may be protracted to all ASCII character codes to take care of specific situation-based requirements:

- Operational ease.
- Minimized error propagation.
- For transmitting as Hard copy.
- Original cover text not required at receiving end.

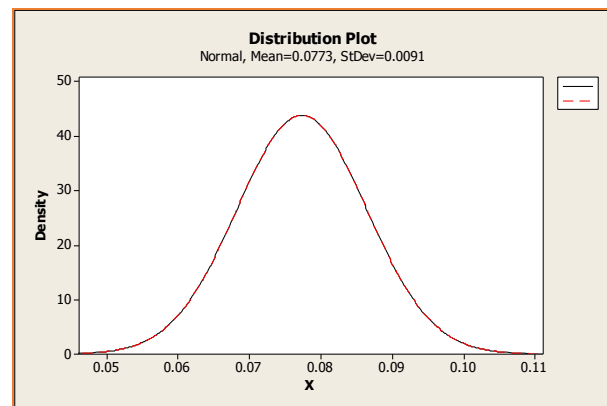


Fig. 7 Probability Distribution Plots for Cover and Stego Object

6. Recommendations

Following are recommended for future enhancement of the proposed scheme:

- Preference be given to use of True Random Number Generator (TRNG).
- With oblivious Stego Key selection mechanism, even originator won't be able to embark on locations used for embedding secret message bits in cover text.

7. Conclusion

This paper presented a secure steganographic cover generation scheme for ASCII text document by first deliberating on traits concerning ASCII character codes such as revelation of underlying data hiding technique e.g. substitution or transposition, by just counting the number of occurrences of English alphabets and contrasting those against their predetermined values in Standard English text. 256-bit Stego key dependent bit embedding is in

adherence to Kerchoff's principle. The statistical characteristics of the stego object (randomness) makes it hard for Warden Wendy not to differentiate it from a cryptogram thereby providing information theoretically secure communication for a time sufficient enough to protect vital information constraint only by her effort towards breaking it. Moreover, the stego object is less prone to errors and can also be transmitted in printed form. Further, availability of cover text at receiving end is not required.

Steganography for Information Security", <http://www.datamark.com.sg/pdf/Steganography.pdf>

- [3] Tacticus, How to survive under siege / Aineias the Tactician, pp. 84-90, 183-193. Clarendon ancient history series, Oxford, England: Clarendon Press, 1990, ISBN 0-19-814744-9, translated with introduction and commentary by David Whitehead.
- [4] Richard Anthony Mollin, Codes - The Guide to Secrecy from Ancient to Modern Times, Chapman & Hall/CRC, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742 ©

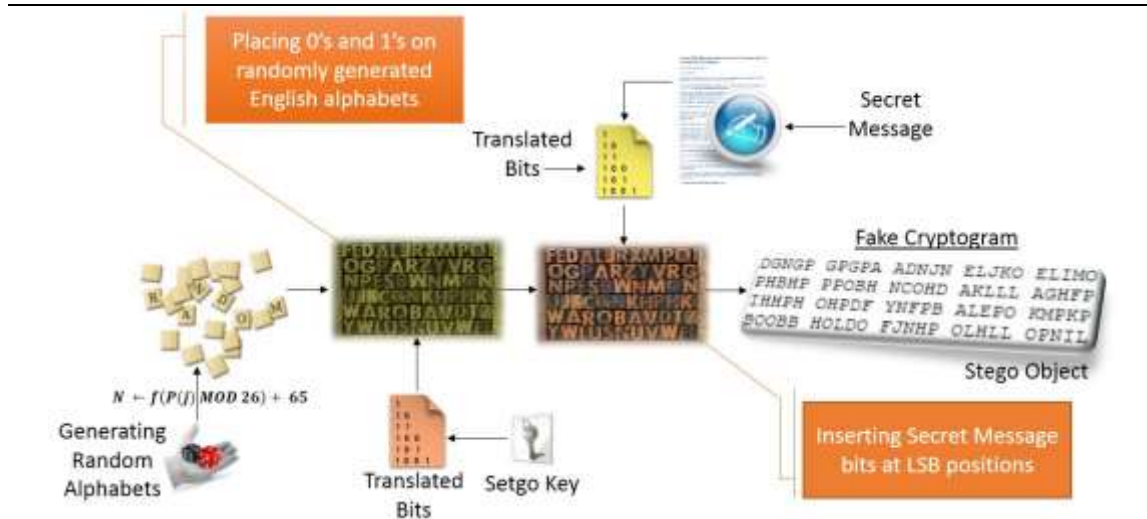


Fig. 8 Secret-Bit Embedding Process

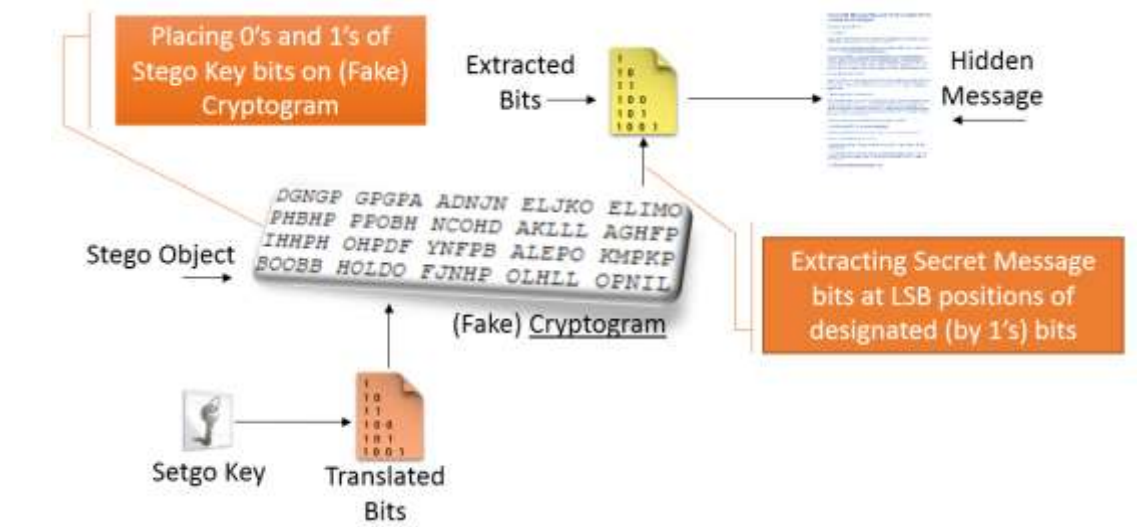


Fig. 9 Hidden-Bit Extraction Process

References

- [1] Krista Bennett, "Linguistic Steganography: Survey, Analysis, And Robustness Concerns For Hiding Information In Text", Cerias Tech Report 2004-13, Center for Education and Research in Information Assurance and Security, Purdue University, West Lafayette, IN 47907-2086
- [2] Anthony T.S. Ho, Siu-Chung Tam, Kok-Beng Neo and Sim-PengThia, Siong-Chai Tan and Lian-Teck Yap, "Digital

2005

- [5] Manly P. Hall, The Secret Teachings of All Ages, Philosophical Research Society Inc., U.S. (June 1989) - 254 Pages, ISBN: 089314830X / ISBN-13: 9780893148300
- [6] Simon Singh, The Code Book: How to make it, Break it, Hack it, Crack it, Delacorte Press an imprint of Random House Children's Books a division of Random House, Inc. 1540 Broadway New York, 10036

- [7] John A. Fraser, The Use of Encrypted, Coded and Secret Communications isan "Ancient Liberty" Protected by the United States Constitution, 2 Va. J.L. & Tech. 2 (Fall 1997) <<http://vjolt.student.virginia.edu>> 1522-1687 / © 1997 Virginia Journal of Law and Technology Association
- [8] Howard, Sir Michael, The Franco-Prussian War: The German Invasion of France, 1870-71, CANTABOOKS, Canterbury, KEN, United Kingdom, ISBN 10: 041502787X / ISBN 13: 9780415027878
- [9] Jacqueline L. Tobin, Raymond G. Dobard, Hidden in Plain View: A Secret Story of Quilts and the Underground Railroad, ISBN 10: 0385497679 / 0-385-49767-9, ISBN 13: 9780385497671, Publisher: Bantam Dell Pub Group, Publication Date: 2000
- [10] Fred Cohen, A Short History of Cryptography, 1995, <http://all.net/edu/curr/ip/Chap2-1.html>
- [11] Bartholomew Lee; Radio Intelligence Developments during World War One and between the Wars; <http://antiqueradios.com/chrs/journal/intelligence.html> [Last Accessed - Dec 05, 2012]
- [12] David Kahn, The Code breakers, The Macmillan Company. New York, NY 1967
- [13] Kenneth Sewell and Jerome Preisler; All Hands Down: The True Story of the Soviet Attack on the USS Scorpion; 277 pages, Published: April 15th 2008 by Simon & Schuster, ISBN # 0-7432-9798-9
- [14] Michael Bates, T-O-R-T-U-R-E, <http://www.batesline.com/archives/2009/02/torture.html>
- [15] Ross Anderson, Stretching the Limits of Steganography, www.cl.cam.ac.uk/~rja14/Papers/stegan.pdf
- [16] Sam Jones; Smithy's code: new twist in Da Vinci drama, The Guardian, Thursday 27 April 2006
- [17] Alan G. Konheim, Computer Security and Cryptography, John Wiley & Sons, Inc., Hoboken, New Jersey copyright 2007
- [18] Dinesh Goyal, Vishal Srivastava, "RDA Algorithm: Symmetric Key Algorithm", International Journal of Information and Communication Technology Research, Volume 2 No. 4, April 2012, ISSN 2223-4985
- [19] Mark S. Granovetter, "The Strength of Weak Ties", American Journal of Sociology, Volume 78, Issue 6 (May, 1973), 1360-1380
- [20] Stefan Katzenbeisser, Fabien A. P. Petitcolas. Information Hiding Techniques for Steganography and digital Watermarking, Artech House, Boston, London, 2000, pp. 26-27.
- [21] R. Chandramouli, Nasir Memon."Analysis of LSB Based Image Steganography Techniques." In proceedings of International Conference on Image Processing, 2001, vol.3 pp. 1019-1022, 2001.
- [22] Johnson, N.F. & Jajodia, S. "Exploring Steganography: Seeing the Unseen." Computer Journal, pp. 26-34, February 1998. <http://www.jjtc.com/pub/r2026.pdf>
- [23] Charles Petzold. The Hidden Language of Computer Hardware and Software. Microsoft Press A Division of Microsoft Corporation One Microsoft Way Redmond: Washington, pp. 177, 1999.
- [24] ASCII, en.wikipedia.org/wiki/ASCII
- [25] Ibrahim A. Al-Kadi. "The Origins of Cryptology: The Arab Contributions." Cryptologia 16(2), pp. 97-126, April 1992.
- [26] Jan Krhovjak, Petr Svenda, VashekMatyas, Random Numbers and Mobile Devices, TU Dresden: Hauptseminar Technischer Datenschutz – 2008, http://www.fi.muni.cz/~xkrhovj/index_en.kg.html
- [27] Codegroup, Five-Letter Codegroup Filter, <http://www.fourmilab.ch/codegroup/> [December 17, 2012]
- [28] H. Beker and F. Piper. Cipher Systems: The Protection of Communications. Northwood Books, London, 1982
- [29] Richard C. Waters. "Cryptology and Data Communications." Internet: http://dspace.mit.edu/bitstream/handle/1721.1/41974/AI_WP_136.pdf?sequence=1, December 1976 [July 12, 2012].
- [30] Shannon, C.E. "Communication Theory of Secrecy Systems." Bell System Technical Journal, pp. 656–715, October 1949.
- [31] [T Hansen - 2006, US Secure Hash Algorithms \(SHA and HMAC-SHA\), http://tools.ietf.org/html/rfc4634](http://tools.ietf.org/html/rfc4634)
- [32] Hamming, Richard W. (1950), "Error detecting and error correcting codes", Bell System Technical Journal 29 (2): 147–160, MR 0035935.
- [33] В.И. Левенштейн (1965). "Двоичные коды с исправлением выпадений, вставок и замещений символов". Доклады Академии Наук СССР 163 (4): 845–8. Appeared in English as: Levenshtein VI (1966). "Binary codes capable of correcting deletions, insertions, and reversals". Soviet Physics Doklady 10: 707–10.
- [34] M.A. Jaro., "Probabilistic linkage of large public health data files" (disc: P687-689). Statistics in Medicine, 1995, pp. 14:491-498
- [35] MiniTab 16. <http://www.facebook.com/Minitab> [June 6, 2012]



KHAN FARHAN RAFAT is a Ph.D. Scholar at International Islamic University, Islamabad – Pakistan. He did his MCS from Gomal University, D.I.K. followed by MS in Telecommunication Engineering from UMT, Lahore – Pakistan. A veteran of information security having an experience of almost about 24 years has worked in varied roles in areas not limited only to programming, evaluation & analysis of Software/Hardware based security modules, and formulating security policies.



Professor Dr. Muhammad Sher is Dean Faculty of Basic and Applied Sciences at International Islamic University, Islamabad – Pakistan. He received B.Sc. degree from Islamia University Bahawalpur and M.Sc. degree from Quaid-e-Azam University, Islamabad, Pakistan. He did Ph.D. in Computer Science and Electrical Engineering from TU Berlin, Germany His area of research is Next Generation Networks Security. An eminent Scholar who has a number of research publications to his credit.

Performance analysis of Internet Protocol Storage Area Network (IP SAN) and its usage in Clustered Database

Akanksha Verma¹, Shrijee Bhattacharjee², Prashant Baste³, Nandan Mirajkar⁴
Sandeep Bhujbal⁵, Nilesh Deshmukh⁶, Kumud Singh⁷

^{1,6}Department of Advanced Networking and Telecommunications
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India

^{2,3,4,5}Department of Advanced Software and Computing Technologies
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India

⁷Systems Department
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India

Abstract

In current social networking world, organizations are demanding high-speed security for client data such as images, audio video files even huge databases. Also no flexibility to connect and share remote devices and servers reduces the performance of service given by organization. To overcome this issues Storage Area Network (SAN) was introduced which is a dedicated storage network that carries I/O traffic only between servers and storage devices. The paper describes requirement of SAN in real world by implementing a SAN using Openfiler, Based on seven cases SAN's performance analysis done with help of Iometer tool and SAN's importance in clustered Database.

Keywords: Storage Area Network (SAN), Iometer, Average I/O Response time, Database, Real Application Cluster (RAC).

1.0 Introduction

The SAN improves the concept of data sharing. Although a typical LAN enables applications and end users to access data held in a central location, the SAN moves that data onto a much faster infrastructure. This allows multiple computers to transfer large files concurrently at rates comparable to locally attached disks over the SAN without adversely affecting the corporate LAN [17]. Openfiler is a opensource Linux operating system which can be used to design a SAN. In case of performance analysis five disks with video data are accessed by two machines and analysis is done for various cases on readings taken by Iometer tool. Oracle RAC has benefits including fault tolerance, security, load balancing, and scalability.

Unfortunately, for many shops, the price of the hardware required for a typical production RAC configuration makes this goal impossible. A small two-node cluster can cost from US\$10,000 to well over US\$20,000. This cost would not even include the heart of a production RAC environment, the shared storage. In most cases, this would be a Storage Area Network (SAN), which generally start at US\$10,000. This paper provides a low-cost alternative to configuring an Oracle RAC 11g Release 2 system. All shared disk storage for Oracle RAC will be based on iSCSI (Internet Small Computer System Interface, an Internet Protocol (IP)-based storage networking standard for linking data storage facilities) using Openfiler running on a third node [9]. Section 2 illustrates about SAN its definition, architecture and its benefits while section 3 puts forward the results obtained for performance analysis. Section 4 describes the importance of SAN in clustered database.

2.0 Storage Area Network (SAN)

2.1 Definition

A **Network** is a collection of computers and devices which are interconnected by communication channels. These channels allow sharing of services, resources and information among it efficiently. **Storage Area Network** (SAN) is a network whose primary purpose is to transfer data between computer systems and storage elements-as defined by Storage Networking

Industry Association (SNIA) [1]. A storage area network (SAN) is a dedicated high performance network to facilitate block-level data access. It carries data between servers (hosts) and storage devices through switches [2][5].

2.2 Architecture

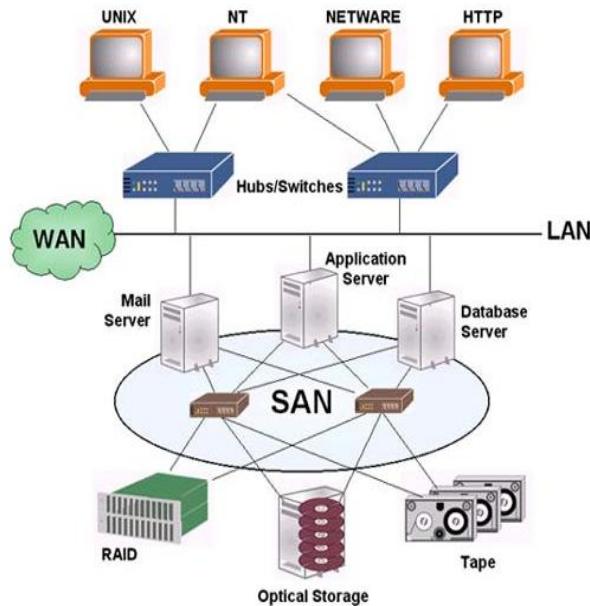


Fig 1: Architecture of Storage Area Network (SAN) [3]

The SANs are used to connect shared storage arrays and tape libraries to multiple servers, and are used by clustered servers for failover. A SAN allows direct, high-speed data transfers between servers and storage devices, potentially in any of the following three ways:

- 1. Server to storage:** This is the traditional model of interaction with storage devices. The advantage is that the same storage device might be accessed serially or concurrently by multiple servers.
- 2. Server to server:** A SAN might be used for high-speed, high-volume communications between servers.
- 3. Storage to storage:** This outboard data movement capability enables data to be moved without server intervention, therefore freeing up server processor cycles for other activities like application processing [1].

The information stored in SAN can be accessed by all servers via Local Area Network (LAN) and Wide Area Network (WAN) so it becomes easy for information accessing. IP SAN uses TCP/IP as its media. The Transmission Control Protocol (TCP) and the Internet Protocol (IP) is part of the backbone of the Internet's suite of communication protocols. The advantage of IP SAN is that when it is utilized, networked storage can

be available any place TCP/IP goes. Internet SCSI (iSCSI) uses the SCSI command set to communicate between the computing devices and storage, via a TCP/IP network. IP SAN uses TCP as a transport mechanism for storage over Ethernet, and iSCSI encapsulates SCSI commands into TCP packets, thus enabling the transport of I/O block data over IP networks [4].

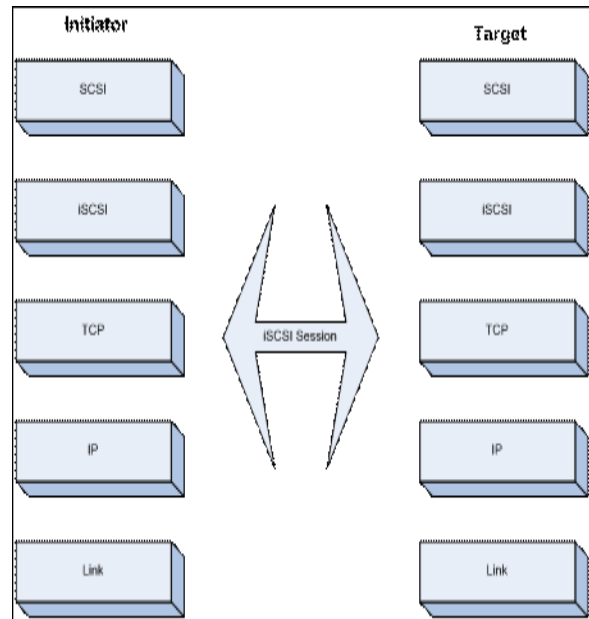


Fig 2: iSCSI and IP SAN Architecture [4]

2.3 Following are benefits of SAN:

1. Removes the distance limits of SCSI-connected disks.
2. Greater performance.
3. Increased disk utilization.
4. Higher availability to storage by use of multiple access paths.
5. Reduced data center rack/floor space.
6. New disaster recovery capabilities.
7. Online recovery.
8. Better staff utilization [18].

3.0 Performance analysis of IP SAN

Performance analysis of IP SAN can be done with the help of **Iometer**. Iometer is both a workload generator (it performs I/O operations in order to stress the system) and a measurement tool (it examines and records the performance of its I/O operations and their impact on the system). It can be configured to emulate the disk or network I/O load of any program or benchmark, or can be used to generate entirely synthetic I/O loads. It can generate and measure loads on single or multiple (networked) systems [16].

Iometer can be used for measurement and characterization of:

1. Performance of disk and network controllers.
2. Bandwidth and latency capabilities of buses.
3. Network throughput to attached drives.
4. Shared bus performance.
5. System-level hard drive performance.
6. System-level network performance.

Following are some parameters with their explanations required to study and perform analysis on data.

Total I/Os per Second: Average number of I/O operations per second, averaged over the length of the test so far.

Total MBs per Second: Average number of Megabytes read and written per second, averaged over the length of the test so far.

Average Latency:

a. **Average I/O Response Time (ms):** Average time between initiation and completion of an I/O operation, averaged over the length of the test so far, in milliseconds.

b. **Average Read Response Time (ms):** Average time between initiation and completion of a read operation.

c. **Average Write Response Time (ms):** Average time between initiation and completion of a write operation.

d. **Average Transaction Time (ms):** Average time between initiation of a request and completion of the corresponding reply. If there are no replies in the access specification, this is the same as Average I/O Response Time.

% CPU Utilization (total): Percentage of processor time spent executing threads other than the Idle thread (in other words, time spent doing useful work). Also known as % Processor Time [16].

While testing various cases some default settings were done in Iometer

1. Maximum disk size 2048000 sectors
2. Access specification : Default
3. update frequency: 10 seconds
4. Run time: 5 minutes

Sample video data was stored in 5 disks created using openfiler and with help of iSCSI initiator all disks were accessed at Machine 1 (M1) and Machine 2 (M2) following is the image for iSCSI status at target.

iSCSI status		
Open sessions for iSCSI target iqn.2006-01.com.openfiler:tan.videosan		
Identifier	Initiator Name	Connections
564050387861568	iqn.1991-05.com.microsoft:ast-tf12-66	View
282575411150912	iqn.1991-05.com.microsoft:ast-tf12-65	View

Fig 3: Screenshot of iSCSI status

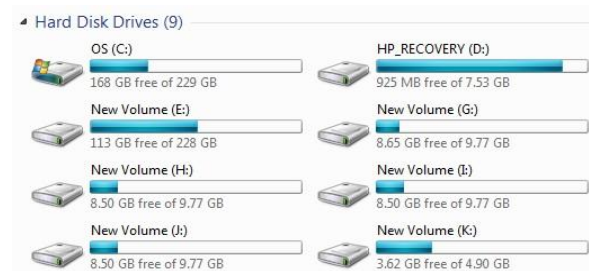
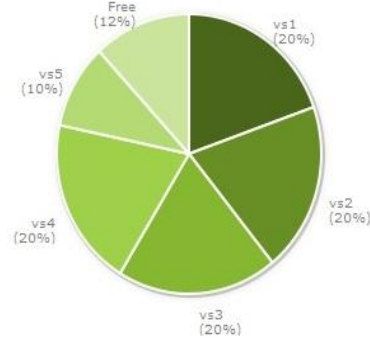


Fig 4: Screenshot of disks available on a machine on connecting to target using iSCSI initiator (G,H,I,J,K)

Volumes in volume group "video_san" (51168 MB)



Volume name	Volume description	Volume size	File system type
vs1	vs1	10016 MB	iSCSI
vs2	vs2	10016 MB	iSCSI
vs3	vs3	10016 MB	iSCSI
vs4	vs4	10016 MB	iSCSI
vs5	vs5	5024 MB	iSCSI
0 MB allocated to snapshots			
6080 MB of free space left			

Fig 5: Screenshot of Volumes in Volume group

Readings were taken for following 7 cases

1. Default case: no video streaming
2. Multiple machines accessing one same video
3. Multiple machines accessing multiple videos (same disk)
4. Multiple machines accessing multiple videos (different disks)
5. At a time one machine accessing one video
6. At a time one machine accessing multiple videos (same disk)
7. At a time one machine accessing multiple videos (different disks)

1. For Local Area Network (LAN)

Sr.No	Cases	Machines	Total I/Os per second	Total MBs per second	Average I/O Response Time (ms)	% CPU Utilization
1	Case1	M1	160.94	0.31	31.0385	4.51%
		M2	163.30	0.32	30.6459	4.85%
2	Case2	M1	158.17	0.31	6.3193	15.77%
		M2	149.84	0.29	6.6703	12.50%
3	Case3	M1	182.81	0.36	5.4676	21.52%
		M2	191.62	0.37	5.2156	14.82%
4	Case4	M1	117.06	0.23	17.0776	20.09%
		M2	119.67	0.23	16.7087	19.72%
5	Case5	M1	431.14	0.84	2.3178	18.29%
		M2	403.18	0.79	2.4778	10.98%
6	Case6	M1	433.95	0.85	2.3024	25.01%
		M2	374.04	0.73	2.6716	23.67%
7	Case7	M1	206.84	0.40	9.6666	20.59%
		M2	196.09	0.38	10.1965	17.49%

Table 1: Readings taken when machines are connected in LAN

For LAN in Case 1 i.e the Default case: no video streaming the average I/O response time for machine 1 M1 is 31.0385 ms and for machine 2 M2 is 30.6459 which are maximum as compared to other cases. For cases 2,3,5,6 where video access is from on disk the response time is less but when video is accessed from different disks the average I/O response time shoots up as seen in cases 4 & 7. Total MBs/sec is high in cases 5 & 6 where at a time one machine is accessing one video and multiple videos from same disk respectively as compared to other cases.

Machine 1: Case 4

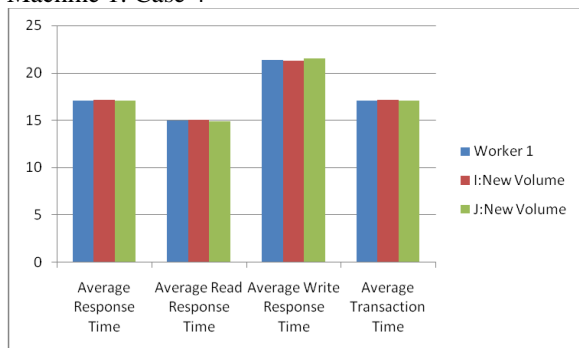


Fig 6: Bar chart for Average Latency in Case4 of Machine 1 (LAN)

Machine 2: Case 4

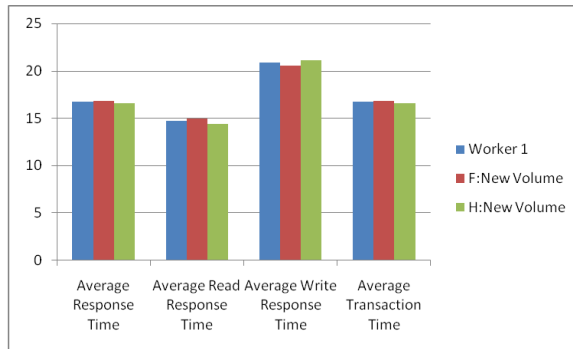


Fig 7: Bar chart for Average Latency in Case4 of Machine 2 (LAN)

2. For Wide Area Network (WAN)

Sr.No	Cases	Machines	Total I/Os per second	Total MBs per second	Average I/O Response Time (ms)	% CPU Utilization
1	Case1	M1	507.35	0.99	9.8485	2.64%
		M2	409.98	0.80	12.1943	6.27%
2	Case2	M1	203.19	0.40	4.9192	13.96%
		M2	218.66	0.43	4.5697	6.76%
3	Case3	M1	205.04	0.40	4.8748	19.76%
		M2	236.81	0.46	4.2194	42.08%
4	Case4	M1	219.77	0.43	9.0973	18.52%
		M2	257.91	0.50	7.7523	29.62%
5	Case5	M1	258.06	0.50	3.8733	16.20%
		M2	293.52	0.57	3.4046	8.00%
6	Case6	M1	232.69	0.45	4.2944	24.01%
		M2	260.18	0.51	3.8404	10.18%
7	Case7	M1	344.87	0.67	5.7970	19.34%
		M2	305.52	0.60	6.5433	10.09%

Table 2: Readings taken when machines are connected in WAN

For WAN the average I/O response time is maximum in Case 1 for both machines. And when video access from different disks is involved i.e Cases 4 & 7 the average I/O response time is high compared to other Cases. In case of WAN the Total MBs/sec is high in Case 1 as compared to other cases.

Machine 1: Case 4



Fig 8: Bar chart for Average Latency in Case4 of Machine 1 (WAN)

Machine 2: Case 4

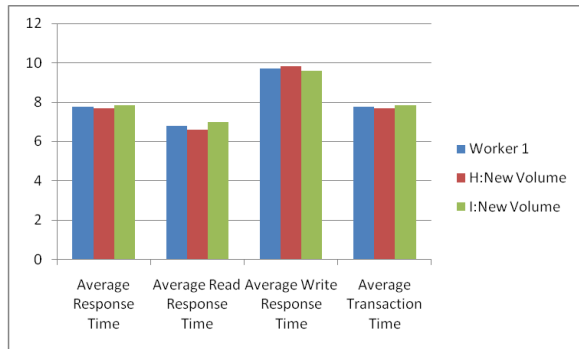


Fig 9: Bar chart for Average Latency in Case4 of Machine 2(WAN)

The Bar charts depicts average write response time in case of LAN or WAN is high as compared to average read response time or average transaction time.

4.0 Oracle Real Application Cluster (RAC)

Oracle’s Real Application Clusters (RAC) supports the transparent deployment of a single database across pools of server, providing fault tolerance from hardware failures or planned outages [7]. In an Oracle RAC environment, two or more computers each with an Oracle Relational Database Management system (RDBMS) instance simultaneously access a single database which makes it possible for an application or user to connect to either computer and have access to a single coordinated set of data. The database consists of a collection of data files, control files, and redo logs located on disk. The instance comprises the collection of Oracle-related memory and operating system processes that run on a computer system [8].

4.1 Usage of SAN in Clustered environment

The Internet Small Computer System Interface (iSCSI) is an Internet Protocol (IP)-based storage networking standard for establishing and managing connections between IP-based storage devices, hosts, and clients. iSCSI is a data transport protocol defined in the SCSI-3 specifications framework and is similar to Fibre Channel in that it is responsible for carrying block-level data over a storage network. In Block-level communication data is transferred between the host and the client in chunks called blocks. Database servers depend on this type of communication [9]. Oracle Real Application Clusters is a shared everything architecture. All servers in the server pool share all storage used for an Oracle RAC database. The type of storage pool used can be network attached storage (NAS), Storage Area Network (SAN), or SCSI disks [10]. We will be using Storage Area Network (SAN) by Openfiler.

Oracle Real Applications Clusters

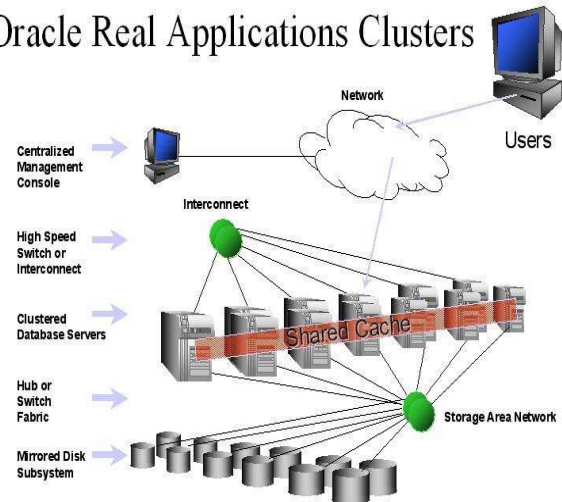


Fig 10: Architecture of Oracle RAC [10]

iSCSI Initiator: An iSCSI initiator is a client device that connects and initiates requests to some service offered by a server (in this case an iSCSI target). The iSCSI initiator software will need to exist on each of the Oracle RAC nodes (racnode1 and racnode2).

iSCSI Target: An iSCSI target is the "server" component of an iSCSI network. It is the storage device that contains the information you want and answers requests from the initiator(s). In this project Openfiler will be the iSCSI target.

4.2 Implementation of Oracle RAC

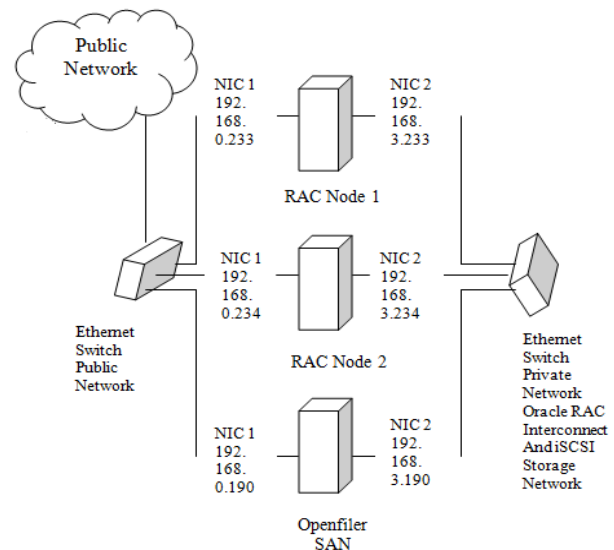


Fig 11: Architecture of Implemented Oracle RAC with IP addresses

In Oracle VM VirtualBox [12] both RAC 1 and RAC 2 machines are built having Operating system Windows server 2008 [11] and following softwares

1. Oracle Grid Infrastructure 11g Release 2
2. Oracle Database 11g Release 2 [13]

```

SQL*Plus: Release 11.2.0.1.0 Production on Thu May 30 16:47:34 2013
Copyright (c) 1982, 2010, Oracle. All rights reserved.
Enter user-name: sys as sysdba
Enter password:
Connected to an idle instance.

SQL> startup
ORACLE instance started.

Total System Global Area 855982000 bytes
Fixed Size 2180544 bytes
Variable Size 654314048 bytes
Database Buffers 197132288 bytes
Redo Buffers 2355200 bytes
Database mounted.
Database opened.
SQL> select * from U$ACTIVE_INSTANCES;

INST_NUMBER
-----
INST_NAME
-----
1
RAC1:orc11
2
RAC2:orc12
SQL>
    
```

Fig 12: Screenshot of Active Instances in RAC

4.3 Time based analysis of retrieve and display data

First we will put 10 lakh records in single machine database (non-RAC environment).

```
create table temp (no int);
```

```

Now using PL-SQL commands to put 10 lakh records
BEGIN
FOR i in 1..1000000 LOOP
INSERT INTO temp(no) VALUES (i);
END LOOP;
COMMIT;
END;
    
```

4.3.1 Time required to put 10 lakh records in single machine database in minutes:seconds

Ta = 1:11.80

4.3.2 Time required to retrieve and display data in single machine database.

```
select * from temp;
```

T1 = 2:30.65
 T2 = 2:30.55
 T3 = 2:28.55
 T4 = 2:29.55
 T5 = 2:29.18

The average time **Tb = 2:29.69**

Now we will put 10 lakh records in tables temp7 from Node 1 and temp8 from Node 2 in RAC database and get retrieve and display results.

4.3.3 Time required to put 10 lakh records in RAC database from node1 and node2 not simultaneously

Nodes	Node 1	Node 2
For Table	temp7	temp8
Time	Tc1 = 3:12.39	Td1 = 3:51.60

Table 3: Time taken to put data in RAC

4.3.4 Time required to put 10 lakh records in RAC database from node1 and node2 simultaneously.

Nodes	Node 1	Node 2
For Table	temp9	temp10
Time	Tc2 = 2:53.70	Td2 = 3:26.97

Table 4: Time taken to put data in RAC simultaneously

As we can compare Ta with Tc's & Td's it proves that time required to put data in RAC environment is more than time required to put data in single machine database. From above readings it can analyzed that time difference between nodes is negligible in RAC environment hence as much as nodes increase performance remains same. In case of production environment with **Fibre channel** speed increases tremendously hence multiple nodes can insert data at same time without any data corruption.

4.3.5 Time required to retrieve and display data in RAC environment

Case 1: RAC environment in LAN (Local Area Network) and nodes accessing same data [14]

	Time	Node 1	Node 2
Not Simultaneous Access of data	T1	8:42.46	7:25.55
	T2	6:48.95	6:55.21
	T3	5:16.46	6:53.76
	T4	5:15.46	6:45.18
	T5	5:01.89	7:01.29
	Average	Tm = 6:13.04	Tn = 7:00.02
Simultaneous Access of data	T1	7:23.27	8:10.73
	T2	7:08.12	7:32.93
	T3	6:57.66	7:18.99
	T4	7:00.74	7:26.98
	T5	7:02.11	7:21.08
	Average	To = 7:06.38	Tp = 7:34.14

Table 5: Time taken to retrieve & display data from RAC (LAN + same data)

Case 2: RAC environment in LAN (Local Area Network) and nodes accessing different data [14]

	Time	Node 1	Node 2
Not Simultaneous Access of data	T1	6:24.81	9:34.29
	T2	5:21.98	6:55.16
	T3	5:27.99	6:28.76
	T4	5:29.12	6:14.91
	T5	5:21.68	6:19.42
	Average	Tq = 5:37.12	Tr = 7:06.51
Simultaneous Access of data	T1	4:33.97	7:25.40
	T2	4:27.17	6:34.78
	T3	4:25.13	7:09.28
	T4	4:24.67	6:56.37
	T5	4:32.90	7:03.11
	Average	Ts = 4:28.77	Tt = 7:01.79

Table 6: Time taken to retrieve & display data from RAC (LAN + different data)

The above 2 cases proves even in case of retrieve and display of data, time delay is minimum, data can be accessed from many servers simultaneously without corruption. The major advantage here is that data is stored in a centralized storage and multiple nodes can perform transactions on that data through network. Hence RAC provides high availability i.e If a node in a server pool fails, the database continues to run on the remaining server in the pool [7].

5. Conclusion

The paper clearly illustrates the performance analysis of SAN considering seven cases and its importance as centralized storage access to all nodes in case of clustered database. The cost and complexity of Fibre Channel has kept SAN deployment out of reach for small and mid-sized businesses until the introduction of Storage over IP (SoIP) SANs based on the iSCSI protocol approved by the Internet Engineering Task Force (IETF) in 2003. In case of security Fibre Channel SANs are traditionally less secure than iSCSI. Fibre Channel SAN can see multiple LUNs on any particular disk, but iSCSI can only deal with a disk target. Consequently, iSCSI authentication is very important, and iSCSI employs advanced authentication methods to establish security, such as Challenge Handshake Authentication Protocol (CHAP). Fibre Channel does not support native

encryption over the wire, but iSCSI can utilize IPsec encryption to protect data in flight [6]. In SAN if all the hosts are allowed to access all the drives the two important problems arise they are disk resource contention and data corruption. To deal with them one can isolate and protect storage devices on a SAN by using zoning and LUN (Logical Unit Number) masking, which allows to dedicate storage devices on the SAN to individual servers. Zoning: Many devices and nodes can be attached to a SAN, When data is stored in a single cloud, or storage entity, it is important to control which hosts have access to specific devices. Zoning implemented at the hardware level, isolates a single server to a group of storage devices or a single storage device, or associate a grouping of multiple servers with one or more storage devices, as required in a server cluster deployment. LUN masking: performed at the storage controller level, allows to define relationships between LUNs and individual servers. Storage controllers usually provide the means for creating LUN-level access controls that allow access to a given LUN by one or more hosts. By providing this access control at the storage controller, the controller itself enforces access policies to the devices. LUN masking provides more detailed security than zoning, because LUNs provide a means for sharing storage at the port level [15]. SAN holds its importance even in this Cloud era and will continue to provide optimum service to end users.

Member Disks

Select Disk	Failure Group	Path	Library	Read/Write Errors	State	Mode	Size (GB)	Used (GB)	Used (%)	Failgroup Type
<input type="checkbox"/> SAN_0000	SAN_0000	\\.\ORCLDISKSANO	SYSTEM	0	NORMAL	✓	9.97	2.33	23.35	REGULAR

Fig 13: Screenshot of disk available in RAC

Disk Group I/O Cumulative Statistics

Member Disks	Average Response Time (ms)	Average Throughput (MB per second)	Total I/O Calls	Reads				Writes			
				Total	Hot	Cold	Errors	Total	Hot	Cold	Errors
Disk Group - SAN	0.16	0.22	128059	103689	0	103289	0	24370	0	20549	0
SAN_0000	0.16	0.22	128059	103689	0	103289	0	24370	0	20549	0

Fig 14: Screenshot of disk statistics

References

- [1] Introduction to Storage Area Networks and System Networking Jon Tate Pall Beck Hector Hugo Ibarra Shanmuganathan Kumaravel Libor Miklas IBM Fifth Edition (November 2012) © Copyright International Business Machines Corporation 2012. All rights reserved.
- [2] http://en.wikipedia.org/wiki/Storage_area_network [last modified 6 June 2013]
- [3] <http://www.allsan.com/sanoverview.php3> [last modified 6 June 2013]
- [4] http://www.tns.com/ip_san.asp [last modified 25 June 2013]
- [5] Information Storage and Management Participant guide volume 1 of 2 student guide EMC2 July 2009
- [6] http://www.cuttedge.com/files/iscsi_vs_fiberchannel_explain.pdf [last modified 24 June 2013]
- [7] Oracle Real Application Cluster , Oracle Data Sheet Copyright © 2010, Oracle and/or its affiliates. All rights Reserved.
- [8] http://en.wikipedia.org/wiki/Oracle_RAC [last modified 6 June 2013]
- [9] <http://www.oracle.com/technetwork/articles/hunter-rac11gr2-iscsi-088677.html> [last modified 6 June 2013]
- [10] An Oracle White Paper November 2010 Oracle Real Application Clusters (RAC) 11g Release 2 Author: Barb Lundhild, Markus Michalewicz Copyright © 2010, Oracle and/or its affiliates. All rights reserved
- [11] <http://www.microsoft.com/en-in/server-cloud/windows-server/2008-r2-trial.aspx> [last modified 6 June 2013]
- [12] <https://www.virtualbox.org/> [last modified 10 June 2013]
- [13] <http://www.oracle.com/technetwork/database/enterprise-edition/downloads/index.html> [last modified 10 June 2013]
- [14] <http://www.csgnetwork.com/timescalc.html> [last modified 10 June 2013]
- [15] [http://technet.microsoft.com/en-us/library/cc758640\(v=ws.10\).aspx](http://technet.microsoft.com/en-us/library/cc758640(v=ws.10).aspx) [last modified 10 June 2013]
- [16] <http://csis.pace.edu/~lombardi/sciences/computer/systems/windows/docs/iometer.pdf> [last modified 10 June 2013]
- [17] <http://www.eetimes.com/design/communications-design/4017917/Storage-Area-Network-Overview> [last modified 23 June 2013]
- [18] http://media.wiley.com/product_data/excerpt/38/04703851/0470385138.pdf [last modified 23 June 2013]

First Author: Ms. **Akanksha Verma**, She is pursuing **M.Tech** in Advanced Information Technology with specialization in Networking and Telecommunications from IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. She completed **B.E** Computer Science and Engineering (CSE) from Rajiv Gandhi Prudyogiki Vishwavidyalaya (RGVP). Her research interest includes Storage Area Network (SAN) and Networking.

Second Author: Ms. **Shrijee Bhattacharjee**, She is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. She completed **B.Tech** Computer Science and Engineering (CSE) from North Eastern Hill University (NEHU). Her research interest includes Cloud Computing, Database Management System, Data Mining and Business Intelligence.

Third Author: Mr. **Prashant Baste**, He is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU–I²IT Centre of Excellence for Advanced Education and Research, Pune, India. He completed **B.E** Computer Science and Engineering from University of Pune. His research interests include databases, data mining and cloud computing.

Fourth Author: Mr. **Nandan Nagarajappa Mirajkar**, He is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. He is also **Teaching Assistant** in Advanced Software and Computing Technologies Department. He completed **B.E** Electronics and Telecommunications from University of Mumbai. His research interests include Cloud computing, Databases and Networking.

Fifth Author: Mr. **Sandeep Bhujbal**, He is **Sr. Research Associate** in Advanced Software and Computing Technologies Department of IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. He completed **M.C.S** from University of Pune. His research interests include Operating systems, Compiler construction, Programming languages and Cloud computing.

Sixth Author: Mr. **Nilesh Deshmukh**, He is **Senior Lecturer** in Advanced Software and Computing Technologies Department of IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. He completed APGP from I²IT and **B.E** Computer Science from Amravati University. His research interests include Storage Area Network (SAN) and Networking.

Seventh Author: Mr. **Kumud Singh**, He is Senior Executive IT Infrastructure Services in IGNOU – I²IT Centre of Excellence for Advanced Education and Research Pune, Maharashtra, India. He completed **BCA** from GIMT Delhi, India. His key interests are in Virtualization and Network security.

Towards an Assessment-oriented Model for External Information System Quality Characterization

Abir ELMIR¹, Badr ELMIR² and Bouchaib BOUNABAT³

AI-Qualsadi Research & Development Team
Ecole Nationale Supérieure d'Informatique et d'Analyses des Systèmes ENSIAS,
Université Mohammed V – Souissi, Rabat Maroc

Abstract

Information System Quality (ISQ) management discipline requires a set of assessment mechanisms to evaluate external quality characteristics that are influenced by the environmental parameters and impacted by the ecosystem factors. The present paper suggests a new assessment oriented model that takes into consideration all facets of each external quality feature. The proposed model, named RatQual, gives a hierarchical categorization for quality. RatQual is designed to quantify dependent-environment qualities by considering internal, external and in use aspects. This model is supported by a tool that automates the assessment process. This tool gives assistance in quality evolution planning and serves for periodical monitoring operations used to enhance and improve information system quality.

Keywords: *Information system quality management, External quality characteristics, Quality assessment, RatQual Model.*

1. Introduction

To support inter organizational collaboration, there is an increasing trend for several information systems to span boundaries between organizations. While information systems quality (ISQ) enhancement is a mean to increase performance and to achieve strategic goals, many investments are done in order to improve quality levels. Studying these systems implies a specific focus on external information system quality (EISQ) characteristics influenced by the environmental parameters and impacted by the ecosystem factors.

Therefore, the domain of ISQ management is a subject to numerous engineering and research works. Many efforts are done in order to propose approaches to ensure and control quality. One of the active branches deals with the characterization and the assessment techniques. Indeed, many models are proposed to describe quality attributes and their relationships. Despite the richness and benefits of these models, they present several disadvantages regarding the consideration of all operational aspects in the characterization and assessment processes.

This work aims to propose a new assessment oriented model designed to characterize EISQ. This model clears the way to have adequate mechanisms in order to assess EISQ characteristics. Such mechanisms take into account internal, external and in use aspects.

This paper is organized as follows. The next section reminds the principal axis of EISQ management discipline. It introduces also ISQ assessment approaches. Section 3 is in relation with the main contribution of this work. It describes the new model proposed to characterize and assess EISQ. This section proposes a categorization for EISQ characteristics based on quality requirement management perspective. It notes the importance to take into consideration internal, external and in use aspects of each external quality characteristic. It enumerates the main functionalities of Quality Monitoring Tool that automates the assessment activities. This is followed by the conclusion in Section 4.

2. Quality characterization and assessment

2.1 External quality of inter organizational systems

Organizations are increasingly concerned with the quality management of their information systems and make continuous investments to enhance their different qualities levels. These investments are justified by the importance of ISQ in order to achieve organizational objectives and to increase performance and profits [1].

Also, there is an increasing trend for several information systems to span boundaries between organizations. Such systems can be used to support collaborations and partnerships among organizations for competitive purposes. Low quality level of inter organizational systems is a potential failure of cooperation and collaboration [2].

Within a collaborative ecosystem, ISQ improvement deals with conceptual, organizational and technical barriers between stakeholders that may belong to different governance subdomains [3].

For this purpose, Studying ISQ implies a specific focus on its external characteristics that are influenced by environmental parameters and impacted by the ecosystem factors. As example of such characteristics, we quote (i) interoperability, (ii) security, (iii) adaptability, (iv) flexibility, (v) horizontal alignment ability.

Many efforts are done in order to propose approaches to ensure, manage and control ISQ characteristics. In this area, one of the active branches deals with the characterization and the assessment techniques [4]. Indeed, many models are proposed to characterize quality attributes and their relationships. Despite the richness and benefits of these models, they present several disadvantages regarding the consideration of all operational aspects in assessment [3, 5].

2.2 Quality models

The domain of information system quality is subject to numerous modeling initiatives. The first ones lead to hierarchical definitions of quality factors composed of characteristics, which lead to evaluations based on metrics [6, 7, 8]. Such models represent specifications of ISQ characteristics. They (i) relate various quality attributes, (ii) identify practices to address them and (iii) describe metrics for measuring or observing them [9].

ISQ models are useful in (a) the requirements identification and their completeness validation processes, (b) the identification of design and testing objectives and the user acceptance criteria and (c) the communication improvement between all information system stakeholders (acquirers, architects, developers, etc.) [9].

Information system engineering researchers and practitioners have suggested many different quality models. The well-adopted ones characterize ISQ as a hierarchical multidimensional system [10]. There are also other models that adopt relational topology [11].

Other models exist and are star-based topology [12] or have Bayesian Belief Network (BBN) topology [13]. In spite of the number of similarities, each proposed model has its own terminology and has a varying number of attributes as illustrated in Table 1 below.

ISO/IEC 9126-1[14] quality model incorporates the findings of previous models. ISO/IEC 9126-1 includes six characteristics (Functionality, Reliability, Usability, Efficiency, Maintainability, And Portability), which are further subdivided into 27 sub-characteristics. For instance, functionality characteristic includes a set of sub-characteristics: security, interoperability, suitability, accuracy, compliance with standards.

Table. 1: ISQ models components

ISQ Model	Topology	First / Second Level	First level Characteristics
Mccall (1976) [6]	H*	Factor / Criteria	Correctness, reliability, efficiency, integrity, usability, maintainability, testability, flexibility, portability, reusability, and interoperability
Boehm (1978) [7]	H*	High level charac.*/ Primitive charac.	Testability, understandability, efficiency, Modifiability, reliability, portability, and human Engineering
Murine (1983)[15]	H*	Factor / Criteria	Correctness, Reliability, Efficiency, Integrity, Reusability, Usability, Maintainability, Testability, Flexibility, Portability, Interoperability, Intraoperability
Bowen (1985)[16]	H*	Factor / Criteria	correctness, Reliability, Efficiency, Usability, Integrity, Maintainability, verifiability, Portability, flexibility, reusability, interoperability, survivability, expandability
Evans and Marciniak (1987)[17]	H*	Factor / Criteria	Mccall (1976) [6] factors + Verifiability, Expandability
FURPS (1987)[18]	H*	Charac./ Sub charac.	Functionality, usability, reliability, performance, supportability.
Gillies (1987)[19]	R*	-	Maintainability, Flexibility, Testability, Portability, Reusability, Interoperability, Correctness, Reliability, Efficiency, Integrity, Usability
Deutsch and Willis (1988)[20]	R*	Factor / Criteria	Clarity, Integrity, Traceability, Reliability
ISO/IEC 9126 (1991)[14]	H*	charac. / Sub charac.	Reliability, maintainability, Portability, usability, functionality, and efficiency
Dromey (1996)[21]	H*	H-level attribute / Sub attribute	Maintainability, Reliability, efficiency, usability, portability, Reusability, functionality
IEEE 1061 (1998)[22]	H*	Factor / Subfactor	Efficiency, functionality , maintainability , portability , reliability, usability
Perry (1987)[11]	R*	-	Correctness, Reliability, Efficiency, Integrity, Usability, Maintainability, Testability, Flexibility, Portability, Reusability, Interoperability
Stefani et al. (2003)[13]	B*	-	Operability, Reliability, Functionality, Efficiency
Khosravi (2004)[12]	S*	-	Usability, Understandability, Learnability., Operability, Flexibility, Reusability, Robustness, Environmental tolerance, error tolerance, failure tolerance, scalability
AOSQUA MO (2009)[23]	H*	charac. / Sub charac.	ISO/IEC 9126 charac.

H* : Hierarchical, R* : Relational, S* : Star based, B* : BBN based
 charac.*: characteristic

In fact, ISQ can be viewed from various perspectives. Several taxonomies have been proposed in this context. In this sense, there are (see Fig 1):

- Many levels of ISQ concern: business, process, service and data level [24].
- Various approaches to establish ISQ: integrated, federated, and unified approach [25].
- Multiple barriers could handicap ISQ establishment: conceptual, organizational and technical barriers [26].
- Different scopes of application: within the same organization, cross independent organizations [27],
- Different ISQ aspects: internal, external and in use aspects [28].

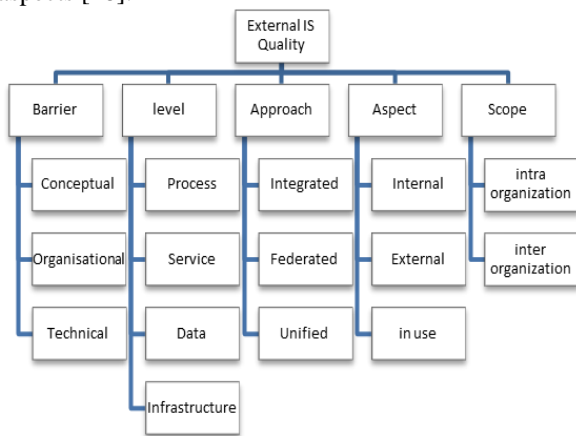


Fig. 1. EISQ taxonomy

2.3 Assessment approaches

In addition to characterization concern, one of the active branches in ISQ management discipline is assessment methods and metrics development. ISQ models, and especially those who belong to “Factor, Criteria, Metric” Family, note the importance of metrics in order to assess quality levels. Also, ISQ management discipline denotes the development of several measurement approaches for the same characteristic. As an illustration, Elmir et al. [29] identifies fourteen different assessment approaches for the single characteristic of interoperability. These approaches have many differentiation aspects like (i) domain of application, (ii) system orientation (qualifying a system or a link between systems), (iii) evaluation instance (a priori, a posteriori), (iv) level character, (v) quantitative or qualitative evaluation, and (vi) coupling with advanced mathematic techniques [29].

In spite of the richness and the advantages of existing assessment approaches, they present several disadvantages regarding the consideration of all operational aspects in assessment. Therefore:

1. The majority of approaches assess “a priori” ISQ degree. Few of them are interested on “a posteriori” aspects. No one of them proposes to take into consideration the both aspects.
2. The existing approaches qualify quality degree of a component or a link between systems. Few of them are able to take in charge the two situations.
3. The majority of approaches is qualitative and describes maturity level with a specific perspective (technical or organizational).
4. Few of the existing approaches combine their results with advanced mathematical techniques such optimization, probability, matrixes (linear algebra), logics or complexity.
5. None of the approaches explicit how prior assessment is used for effective implementation of planned state. Indeed, the measurement process stops when the ISQ degree is calculated. Using this level thereafter as explicit parameter improvement is supported by no one of the existing approaches.

3. RatQual Model: an assessment-oriented Quality model

RatQual (for Ratio of Quality) is an assessment oriented model that proposes an innovative three axis hierarchical classification of 17 EISQ characteristics. The result of the assessment approach is a ratio metric enabling the measurement of specific EISQ characteristic by taking into account three main operational aspects: internal, external and in use ones.

3.1 EISQ RatQual categorization

This work proposes a classification of EISQ characteristics using a requirements engineering perspective. Quality requirements engineering is a discipline interested into formal quality requirements definition and change management.

Requirements engineering applied to quality management area and specifically to EISQ extent implies a three axis categorization. The first axe is about the functional ISQ identification. The two other axes are related to change requests issues. Indeed, the second axe is more interested into context dependent adaptation requests. The third axe is more sensitive to requests evolution over time.

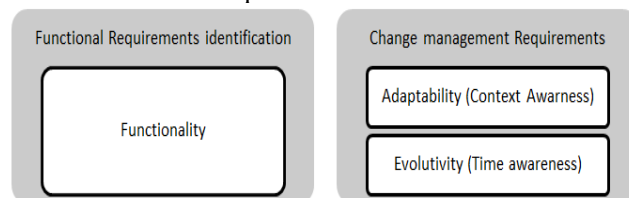


Fig. 2. Requirement perspective in EISQ Classification

Indeed, the first class of characteristics is Functionality. This class refers to the essential purpose of the involved information systems and their components. Functionality characteristics are mainly recognized in the requirements identification stage. This class contains various features among which interoperability, security, compliance and inter-alignment ability.

The other classes we propose are related to quality requirements linked to system change management. Change requests can be classified into two main categories: (i) “Adaptability category” including context dependent change requests, and (ii) “evolutivity category” time dependent change requests.

The former EISQ category entitled “Adaptability” includes Portability, Coexistence, Replace ability, Flexibility and Variability. The latter category named “Evolutivity” encloses characteristics like Changeability, Maintainability, Stability, Testability, Customizing Ability and Extensibility.

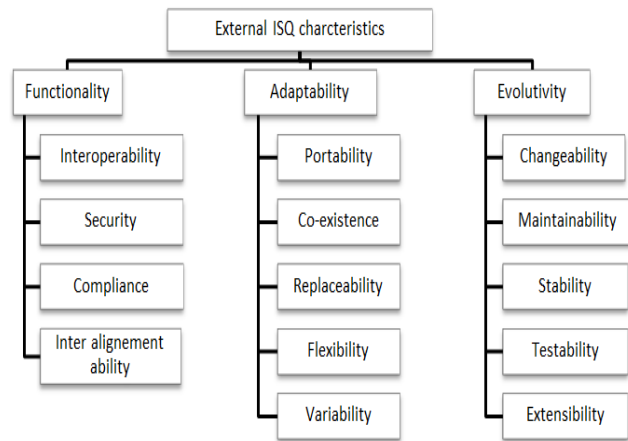


Fig. 3. EISQ characteristics classification

3.2 EISQ appraisal aspects

In terms of appraisal dimensions, ISQ characteristics have a priori and a posteriori assessment details. A posteriori assessment is related to in use operational performance appraised after effort implementation of the desired quality. A priori assessment, for its part, takes into consideration two features: (i) quality potentiality and (ii) quality implementation compatibility.

For this purpose, the proposed model identifies three essential appraisal aspects:

- Quality potentiality: it is an «internal feature» of a system that reflects its characteristic preparation within a collaborative context. This involves identifying a set of requirements that have an impact on interaction

characteristic capacity with partner’s systems without necessarily having concrete information on them. The objective is to foster quality readiness and preparation by eliminating barriers that may reduce the quality degree.

- Quality implementation compatibility: it represents an «external feature». In fact, enhancing the characteristic ability of two support systems is ensured through an engineering process aiming to establish inter organizational collaboration between them and also respond to the desired characteristic requirements.

- Quality performance: the third aspect characterizes the «quality in use». It focuses on monitoring operational performance. It consists of an assessment of the communication infrastructure availability, and the supporting system in general.

3.3 RatQual Model

RatQual is an assessment oriented model intended to describe external characteristics that are influenced by environmental. RatQual aims to evaluate EISQ quality characteristics using a ratio metric.

RatQual metric aggregates a set of sub metrics that asses complementary aspects. These aspects include “a priori” and “a posteriori” aspects. A priori aspects consist of internal aspects in one hand and external aspects on another hand.

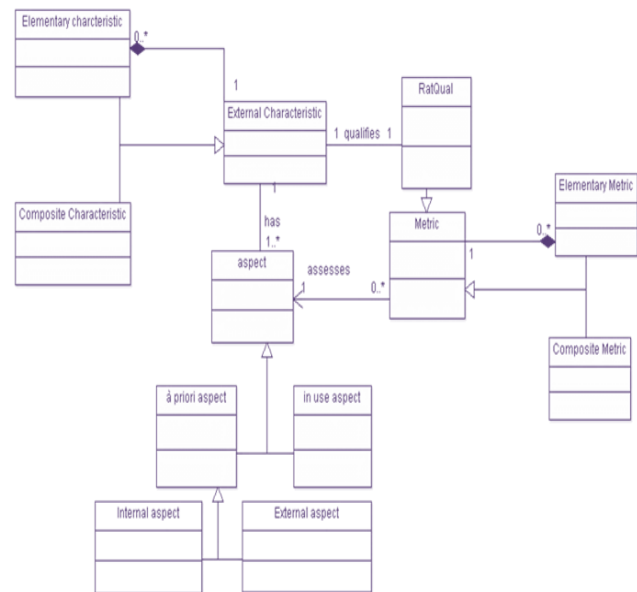


Fig.4. RatQual characterization and assessment model

4. RatQual Assessment Approach

RatQual is a five steps appraisal approach. These steps are as follows (see Figure 5):

1. Delineating the scope of the study.
2. Quantifying the internal aspect: quality characteristic potentiality.
3. Calculating the external aspect: Quality implementation effort.
4. Evaluating the in use aspect: operational quality performance.
5. Aggregating the EISQ RatQual degree based on an adequate aggregation technique.

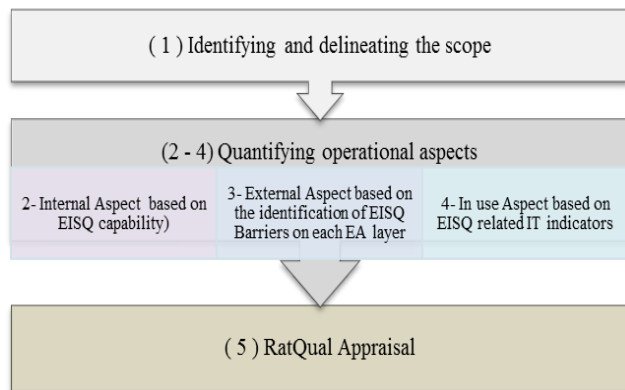


Fig.5. RatQual assessment approach

4.1 Scope delineation

Assessing an external quality characteristic degree of a system requires the knowledge of its ecosystem. In practical terms, the study focuses on a macro business process consisting of a set of sub auto-mated processes among independent business entities. These sub processes are linked together via several interfaces identified in advance. In this case, the preliminary phase consists of identifying the context of the studied automated business process then lists its underlying automated processes. This step includes identifying:

- Organizations involved in the cooperation.
- Sub process within each entity in order to study the compatibility degree of stakeholders in order to enhance a specific EISQ.
- Information systems that support automated business processes within each organization.
- Application services that enables sub processes collaboration.

4.2 Internal aspect: Quality potentiality

The calculation of the potential for quality characteristic within the k_{th} organization « QP_k » requires the adoption of one of the quality maturity models. The organization is classified then on one of the five levels noted QMML (for Quality maturity model level). Table 2 below illustrates a EISQM inventory.

Table. 2. Information System Quality Maturity Models

Quality charac.	Quality maturity model
Functionality	Functionality maturity model integration (FMMI) [30,31]
Interoperability	Interoperability Maturity Model (IMM), Enterprise IMM (EIMM), Organizational IMM (OIMM), Level of Information System Interoperability (LISI) [3, 5, 29]
Security	Information Security Maturity Model (ISMM)[32]
Compliance	Governance Compliance Maturity Model (GoCoMM)[33]
Inter alignment ability	Inter alignment ability maturity model (IAMM) [34]
Adaptability	Quality maturity model (QMM) [35], Adaptability maturity model Integration (AMMI) [30]
Portability	Portability maturity model Integtraion (PMMI) [30]
Co-existence	QMM [35]
Replace ability	QMM[35]
Flexibility	Flexibility maturity model (FMM)[36]
Variability	QMM [35]
Evolutivity	QMM [35]
Changeability	QMM [35]
Maintainability	Architecture Maintainability Maturity Model (AM3) [37], QMM [35]
Stability	QMM [35]
Testability	Testability maturity model (TMM)[38]
Extensibility	QMM [35]

To identify the potential degree of a specific quality characteristic, we propose then the following mapping (See Table 3):

Table. 3. Quantification of quality maturity

Maturity Level (QMML)	Potentiality quantification
1	0.2
2	0.4
3	0.6
4	0.8
5	1

Within each organization, the potential is calculated using the following equation

$$QP_k = 0.2 * QMML_k \tag{1}$$

The final characteristic potentiality is given by Equation 2 below:

$$QP = \min(QP_k) \quad (2)$$

4.3 External aspect: Quality compatibility

To assess the external aspect degree, the present work uses a compatibility matrix [3, 5, 29].

The compatibility matrix, as presented in Table 3, consists of a combination of the “quality levels perspective” and “quality barriers perspective” depicted in Fig.1 of section 2.B. In practical terms, we enumerate conceptual, technical and organizational barriers in the different layers of collaboration concern: process, service, data and infrastructure.

By noting the elementary degree of interoperation compatibility « dc_{ij} » (i takes values from 1..4, and j takes values from 1..6).

Table. 3. Interoperation compatibility

	Conceptual		Organizational		Technology	
	Syntac tic	Seman tic	Respo nsibilit ies	Organi zation	Platfor m	Comm unicati on
Process	dc ₁₁	dc ₁₂	dc ₁₃	dc ₁₄	dc ₁₅	dc ₁₆
Service	dc ₂₁	dc ₂₂	dc ₂₃	dc ₂₄	dc ₂₅	dc ₂₆
Data	dc ₃₁	dc ₃₂	dc ₃₃	dc ₃₄	dc ₃₅	dc ₃₆
Infrastru cture	dc ₄₁	dc ₄₂	dc ₄₃	dc ₄₄	dc ₄₅	dc ₄₆

Therefore, if the criteria in an area marked satisfaction the value 0 is assigned to dc_{ij} ; otherwise if a lot of incompatibilities are met, the value 1 is assigned to dc_{ij} .

The degree of compatibility «DC» is given as follows:

$$DC = 1 - \sum (dc_{ij} / 24) \quad (3)$$

4.4 In use aspect: operating performance

By Denoting:

«DS» the overall availability rate of application servers.

«QoS» service quality of different networks used for interacting components communication. QoS is represented mainly by the overall availability of networks.

«TS» end users satisfaction level about interoperation.

Given the cumulative nature of these three rates, the evaluation of operational performance is given by the geometric mean [31] as the following equation (See Equation 4):

$$PO = \sqrt[3]{(DS * QoS * TS)} \quad (4)$$

4.5 RatQual aggregation

The final calculation of RatQual (for ratio of Quality) is by aggregating the three previous indicators using a function f defined in $[0,1]^3 \rightarrow [0,1]$ (See Equation 5)

$$RatQual = f(PQ, DC, PO) \quad (5)$$

Given the independent nature of these three indicators, we opt for the arithmetic mean [31] as follows (See Equation 6):

$$RatQual = (PQ + DC + PO) / 3 \quad (6)$$

In case we have elements for pondering each one of these three indicators with different weights (w_1, w_2, w_3); we choose the weighted arithmetic mean.

$$RatQual = (w_1 * PQ + w_2 * DC + w_3 * PO) / (w_1 + w_2 + w_3) \quad (6)$$

4.6 Quality monitoring Tool (QMT)

The Quality monitoring tool (IMT) automates the RatQual assessment approach. It includes three principal modules. The first one is dedicated to EISQ characteristic assessment at a specific period. The second one proposes a viable scheme to reach a planned Quality degree. The third module includes a set of reporting views designed to enable periodical quality monitoring activities.

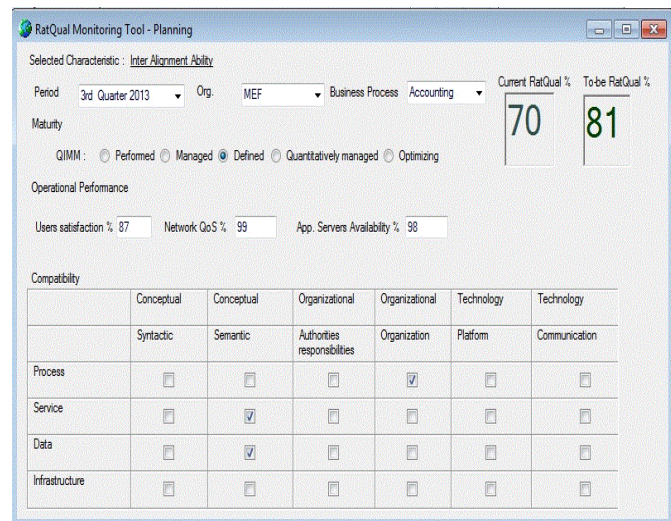


Fig.5. Quality monitoring tool (QMT)

Indeed, QMT has the capacity to track periodically the evolution of quality degree. It gives the possibility to propose a scenario to reach a planned degree of quality characteristic. For instance, in the example shown on Fig. 5, we plan to increase the “inter alignment ability” ratio from an “As-is” degree to a “To-be” one. QMT proposes to (i) improve horizontal alignment maturity to reach the third stage, (ii) optimize the availability of involved application servers, (iii) better meet end users expectations and (iv) to resolve semantic incompatibilities.

4. Conclusions

To operate effectively, organizations are encouraged to enter into close interaction with all their partners. Inter-organizational collaboration is a strategic issue. Quality assurance in this context is very important. For this purpose, this work proposes a novel assessment oriented model for context-dependent quality. This model, named RatQual, takes into account conceptual, organizational and technical considerations and gives importance to architectural elements.

The proposed model here serves to characterize information system external qualities that are influenced by environmental parameters. RatQual considers internal, external and in use aspects. It combines à priori evaluation elements within the design phase of interconnection setup and à posteriori evaluation aspects considering the performance degree of collaboration.

- [1] T. Ravichandran, C. Lertwongsatien, (2005) "Effect of information systems resources and capabilities on firm performance: a resource-based perspective", *Journal of Management Information Systems* 21 (4), pp. 237–276.
- [2] G. Barczak, F. Sultan, EJ. Hultink, (2007) "Determinants of IT usage and new product performance". *Journal of Product Innovation Management*, 24 pp. 600-13.
- [3] B. Elmir, B. Bounabat, (2010) "E-Service Interoperability Measurement within Business Collaboration Networks". *Proceedings of MICS'2010, International Conference on Models of Information and Communication Systems*, Rabat,
- [4] K. Mordal, N. Anquetil, J. Laval, A. Serebrenik, B. Vasilescu, S. Ducasse, (2012) "Software quality metrics aggregation in industry", *Journal of Software: Evolution and Process*,
- [5] B. Elmir, B. Bounabat, (2010) "Integrated Public E-Services Interoperability Assessment" , *International Journal of information Science and Management*, Special issue, October 2010, pp. 1-12.
- [6] J. A. McCall, P. K. Richards, and G. F. Walters, (1977) "Factors in Software Quality", *Nat'l Tech. Information Service*, no. Vol. 1, 2 and 3.
- [7] B.; Boehm, J.R. Brown, H. Kaspar, G.J. MacLeod, M.J. Merritt, (1978), "Characteristics of Software Quality", North-Holland.
- [8] A. Davis, S. Overmyer, K. Jordan, J. Caruso, F. Dandashi, A. Dinh, G. Kincaid, G. Ledebor, P. Reynolds, P. Sitaram, A. Ta, M. Theofanos (1993) "Identifying and measuring quality in a software requirements specification", In *Proceedings First International In Software Metrics Symposium* , pp. 141-152.
- [9] S. Albin, (2003): *The Art of Software Architecture: Design Methods and Techniques*, John Wiley & Sons, March 2003.
- [10] A.C. Stylianou, R. L. Kumar, "An integrative framework for IS quality management", *Communications of the ACM*, Vol. 43(9), pp. 99-104, 2000.
- [11] W. PERRY, (1987), "Effective methods for EDI quality assurance", Prentice-Hall, New Jersey, USA.
- [12] K. Khosravi, Y. Gueheneuc, "A Quality Model for Design Patterns", Technical Report 1249, University of Montreal, 2004.
- [13] A. Stefani, M. Xenos, D. Stavrinoudis, (2003) "Modelling E-Commerce Systems' Quality with Belief Networks", *International Symposium on Virtual Environments, Human-Computer Interfaces, and Measurement Systems*, Switzerland,
- [14] ISO/IEC, ISO/IEC IS 9126, (1991) "Software Product Evaluation - Quality Characteristics and Guidelines for Their Use", *International Organization for Standardization*, Geneva, Switzerland.
- [15] G. E. Murine, (1983) "Improving Management Visibility Through the Use of Software Quality Metrics", *Proceedings from IEEE Computer Society's Seventh International Computer Software & Application Conference*. New York, N.Y.: Institute of Electrical and Electronic Engineers, Inc.
- [16] T. P. Bowen, G. B. Wagle, Tsai, J. T. (1985) "Specification of software quality attributes". *Tech. Rep. RADC-TR-85-37*, Rome Air Development Center.
- [17] M. W. Evans, J. Marciniak, (1987). "Software quality assurance and management", New York, USA: John Wiley & Sons.
- [18] R. Grady, D. Caswell (1987). "Software Metrics: Establishing a Company-wide Program". Prentice Hall. pp. 159.
- [19] A. Gillies, *Software Quality: Theory and Management* (2nd Edition), Internation. Thomson Computer Press, 1997.
- [20] M. Deutsch, R. Willis, 1988. "Software Quality Engineering - A Total Technical and Management Approach", Prentice Hall.
- [21] R.G. Dromey, (1995) "A model for software product quality", *IEEE Transactions on Software Engineering* 21 (2) 146–162.
- [22] IEEE "1061-1998: IEEE Standard for Software Quality Metrics Methodology". IEEE, Los Alamitos (1998)
- [23] K. Avadesh, P.S. Grover, and K. Rajesh (2009). "A quantitative evaluation of Aspect-oriented software quality model (AOSQUAMO) ". *ACM SIGSOFT SEN* 34, 5 (Sep. 2009), 1-9
- [24] L. Guijaro, (2007) "Frameworks for fostering cross-agency interoperability in e-government". *Electronic government: information, and transformation, within the advances in management information system (AMIS) series (ME Sharp)*
- [25] M. Missikoff, F. Schiappelli, F. D'Antonio, (2004) "State of the Art and State of the Practice Including Initial Possible Research Orientations". *Deliverable D8.1. European Commission, INTEROP Network of Excellence*
- [26] W. Y. Arms, D. Hillmann, C. Lagoze, D. Krafft, R. Marisa, J. Saylor, et al., (2002) "A Spectrum of Interoperability: The Site for Science Prototype for the NSDL", *D-Lib Magazine*, 8 (1), January
- [27] W. Guédria, Y. Naudet, D. Chen, (2008) "Interoperability maturity models - survey and comparison", *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg. Vol. 5872/2009, pp. 216–225
- [28] ISO (2001), *International Organization for Standardization: "ISO9126-1:2001, Software Engineering – Product Quality", Part 1: Quality model*
- [29] B. ELMIR, (2012), "Evaluation et optimisation de l'interopérabilité des architectures d'entreprise orientées services, Application aux architectures d'entreprise

gouvernementales", Thèse de doctorat en Informatique de l'ENSIAS - Université Mohamed V, Maroc.

- [30] R. Al Qutaish, (2007). "SPQMM: Software Product Quality Maturity Model using ISO/IEEE standards, metrology, and sigma concepts ", Ecole De Technologie Supérieure (doctoral dissertation, UNIVERSITE DU QUEBEC).
- [31] J. Herbsleb, D. Zubrow, D. Goldenson, W. Hayes, M. Paulk, (1997). "Software quality and the capability maturity model". *Communications of the ACM*, 40(6), 30-40.
- [32] M. F. Saleh, (2011). "Information Security Maturity Model". *International Journal of Computer Science and Security (IJCSS)*, 5(3), 21.
- [33] G. Gheorghe, F. Massacci, S. Neuhaus, A. Pretschner, (2009). "GoCoMM: a governance and compliance maturity model". In *Proceedings of the first ACM workshop on Information security governance* (pp. 33-38). ACM.
- [34] R. G. Santana Tapia, M. Daneva, P. A. T. van Eck, (2007). "Developing an inter-enterprise alignment maturity model: Research challenges and solutions".
- [35] K. C. Wang, J. M. Gao, Z. Y. Gao, J. Q. LIU (2007). "Information quality maturity model". *Computer Integrated Manufacturing Systems*, Beijing, 13(2), 282.
- [36] S. Wadhwa, K. S. Rao, (2002). "Framework for a flexibility maturity model". *Global Journal of Flexible Systems Management*, 3(2), 45-55.
- [37] C. Rathfelder, H. Groenda, (2008). "Towards an Architecture Maintainability Maturity Model (AM3)". *Softwaretechnik-Trends*, 28(4), 3-7.
- [38] I. Burnstein, T. S. AH, G. Saxena, R. Grom, (1999). "A testing maturity model for software test process assessment and improvement", *Software Quality Professional*, 1(4).

A. ELMIR "Ph.D. candidate" at ENSIAS (National Higher School for Computer Science and System analysis) Rabat, holder of a master degree from ENIM Rabat and a mastere diploma from INSA Lyon France (2011). Her research focuses on multi objective optimization and optimal control of information system quality within collaborative networks. She is an integration architect on a private Financial Holding (Banking, Insurance). She also runs the Solutions support activity of this holding.

B. ELMIR He received a Ph.D. degree (2012), an Extended Higher Studies Diploma (2006) and a Software engineer degree (2002) from ENSIAS, (National Higher School for Computer Science and System analysis), Rabat. His research currently focuses on interoperability optimization on public administration. He is an integration architect on the Ministry of Economy and Finance of Morocco since 2002. He also oversees information system quality assurance activity in this department.

B. BOUNABAT Ph.D. in Computer Sciences. Professor in ENSIAS, (National Higher School for Computer Science and System analysis), Rabat, Morocco. International Expert in ICT Strategies and E-Government to several international organizations, Member of the board of Internet Society - Moroccan Chapter.

Towards a better control of the infectious risks associated with health care activities

Benzouai Messaoud¹, Smadi Hacene¹, Kassah-Laouar Ahmed², Bahmani Younes¹ and Bellala-Djamel³

¹ Université Hadj Lakhdar, Laboratoire d'Automatique et de Productique,
Département de Génie Industriel, Batna /Algérie

² Centre Hospitalo-universitaire, Faculté de médecine,
Laboratoire central de microbiologie, Batna / Algérie

³ Université Hadj Lakhdar, Institut d'Hygiène et Sécurité Batna / Algérie

Abstract

The infection associated with health care is a major cause of morbidity and mortality in health establishment. For a dialyzed renal failure, it would be responsible of about 15 % of the deaths. To answer this issue, health establishments use a classical approach, which rests essentially on risk management practices, which enters within the scope of a global analysis of risk management. These practices are built on the anticipation principle; they are based on the identification and the control of the possible risks, of which it is practically impossible to guarantee the exhaustiveness of the risky situations, so it is not possible to leave these anticipation practices at the present time. This is why, for better a control of the infectious risks associated with health care, it is necessary to consolidate more and more this preventive approach and thus to minimize the residual risks. For that purpose an approach is proposed, it is based on a continuous primary prevention strategy, which constitutes a first layer of the protection and is supported by some control and monitoring mechanisms of the critical risks. This approach is illustrated via a real case of study. It is carried out at the hemodialysis center at the teaching hospital, CHU-Batna Algeria.

Keywords: *infectious risk, hemodialysis, VHB, VHC, VIH Bacteremia.*

1. Introduction

The control of the infectious risks, particularly those associated with health care is a major and permanent concern in the health establishments. Indeed the infections associated with health care can have more or less great consequences, going from of a simple pain of the patient and/or the personnel of the establishment, making him faded temporarily the quality of life, until a serious pathology, which leads to an extension in the stay duration, in a long-term disability, and even in a death. Furthermore, additional financial expenses for the patient and the health system are to be taken into account.

In order to answer the caused issue, health establishments use a traditional approach, which rests primarily on the practices of risk management coming from industrial context. This practice enters within the framework of a total analysis of risk management. It is based on the principle of anticipation, which consists of the a priori identification of all the situations with possible risk. For the situation control, some solutions are required like by the design of prevention and protection barriers and by practicing risk analysis. The question is about the control of technological failures, and of the human and organizational errors [1]. However, the literature generally imputes more than 80% of the leading causes to the organization especially when the encountered problems are in the interfaces between the actors or the processes [2].

It is thus practically impossible to guarantee the exhaustiveness of the risky situations, and to admit the assumption that no technical, human, circumstantial or environmental failure will come to disturb the activity progression [3]. In addition to that, the methods of risk analysis do not take into account the following problems: the common mode failures and the non-coherent systems operation [4].

So this traditional approach does not always guarantee the quality of the health care and much less the safety of the patients and the personnel of the health establishment.

Then, the control of the infectious risks requires the implementation of an approach, which stresses upon mechanisms that consolidate more and more the preventive approach and thus minimize the residual risks omitted by the traditional approach. The suggested approach is inspired by the model [5], which is synthesized in figure.1, is articulated around a set of actions allowing a permanent improving of the health care quality, and so minimizing the infectious risks inherent to any health care activity.

This approach rests on a strategy of continuous primary prevention, which represents a first layer of protection continuously consolidated by some mechanisms of identification, control and monitoring of the critical risks.

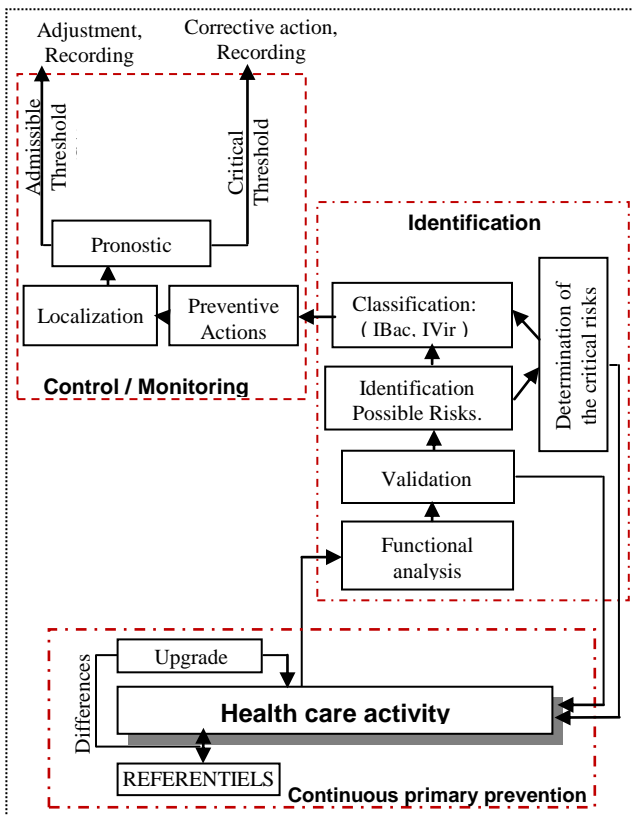


Fig 1. Synoptic diagram of the suggested approach.

The suggested approach is illustrated by a real case of study, representing infectious risks associated to health care. The activity of the chronic renal failure treatment by extra-renal purification (ERP) "Hemodialysis" is an interesting model. The site of action is the Hemodialysis center at the teaching hospital, CHU-Batna Algeria, where several studies in infectious risks are successfully carried out [6, 7, 8, 9].

The topic of the infectious risk in hemodialysis is retained because of the greatest fragility of these patients in final chronic renal failure, regularly exposed to invasive procedures that can put them in contact with various bacterial and viral infectious agents.

2. Field of activity

The hemodialysis is a blood purification technique; it requires the installation of an extra-corporal circulation:

the blood, taken at the arterial level, passes through a membrane (dialysis machine) in order to be filtered by exchange between blood and dialysate. Once purified, it turns over in the patient via the venous circuit. Dialysate circulates and is continuously renewed [10]. This method allows, therefore, to replace the defective renal function, either permanently (chronic hemodialysis, object of our study), or transiently, in waiting for recovering the renal function (grave hemodialysis).

The hemodialysis is an invasive act and requires a repetitive vascular access, on either native arteriovenous prosthetic fistula, or temporary or permanent central catheters venous. Any hemodialysis session involves the risk of transmission of a pathogenic micro-organism on each level of the purification process: dialysis water, concentrated solutions, generator, lines and vascular accesses [11], and also at each stage of the intervention chains, since the arrival of the patient in the dialysis center until his exit.

In hemodialysis, the infectious risk is omnipresent, because of the complexity and the technicality of the health care. This risk concerns the patients, often immunocompromised, but also the health professionals themselves because of the many circumstances of exposure to the biological fluids met during their activity. The infectious risks related to the accidents of exposure to blood (AEB) of the health professionals are not discussed in this paper.

The patients reached of chronic renal failure are abnormally prone to the infections. These infections represent a major cause of morbidity and mortality at the dialyzed renal failure person would be responsible of about 15% for the deaths [12].

The principal infections observed among the hemodialyzed are: 1) The bacterial infections (IBac) which can be of any type but two of them are of prevalent importance in terms of morbidity and mortality and are subject to firstly epidemiologic monitoring. They are vascular access infections (IAV) whose principal micro-organisms in question are in majority Cocci with positive Gram and the bacilli with negative Gram and the bacteremia (BAC) whose principal germs in question are Cocci with positive Gram and enterobacteriaceae. (IAV) and (BAC) are regarded as indicators of the quality of the care [13]. 2) The viral infections (IVir) which are clinically represented by three hematogen viruses that are the virus of hepatitis B (VHB), hepatitis C (VHC) and the virus of the human immunodeficiency (VIH) [14,15,16]. 3) Infections related to the accidents of blood exposure (AEB) altering the health professionals and the infections related to AEB are not approached in this paper.

Like any patient, the hemodialyzed is exposed, on the one hand, with the common infection risks related to any health care activity [17] and on the other hand with the infectious risks specific to the extra-renal purification techniques as well as the working in all the stages of activity like: Contamination of the generator, of the extra-corporal circuit, of the vascular access and the environment. The infectious risks related to the water treatments before the generator are not approached in this paper.

Thus the principal transmission mechanisms of the infections associated with health care in hemodialysis were described namely: endogenous and exogenous mechanisms of contamination, mainly invasive acts, foreign materials, manual contacts and contamination by air.

Through the listed transmission mechanisms, a diagram of as a basis flows is used to identify the various situations at the risk; taking into account on the one hand the conditions and specificities of the hemodialysis center of the CHU of Batna and on the other hand the interactions between the various processes of the hemodialysis activity as well as the various actors.

3. Methods and results

The activity of the chronic renal failure (CRF) treatment by the ERP practice 'our case study' is essentially characterized by a heavy center of hemodialysis, established within the CHU of Batna town, thus allowing the full-time hospitalization of the patient in other service beds. The center comprises 19 stations of the CRF treatment with the trademark "fresenius 4008B and 4008C", of which a rescue station. With a personnel made up of two nephrologists (a head of center and a medical supervisor), the center functions uninterruptedly (except Friday) in 03 sessions per day. For each session, the team is composed of a doctor nephrologist and 7 to 8 nursing graduate status (NGS) including a man of maintenance.

The study was carried out by a multidisciplinary group of the establishment which is sometimes organized in two subgroups working separately on the basis of the same stand, leading to a single result, and this after rapprochement. This is so, with the aim of setting up the maximum reliability to the results.

3.1 Continuous primary prevention

The environment can play a significant role to indirectly facilitate the micro-organisms transmission and to increase the risk of infections.

This phase thus consists in obtaining a favourable environment to the practice of a healthy care activity. It aims at reducing the design intrinsic measurements or upgrade the contamination sources generated by the infrastructure, the logistics and the technical operating conditions of this activity. It is regarded as an essential base to the development of the suggested approach. This measurement allows reducing the primary risk and is regarded as the first layer of protection [18].

Carried out according to lawful referential requirements which specify the existing dialysis methods [19] and detail the logistic characteristics of realization and operation of each one of them [20, 21] as well as the referential of normative requirements, translated by the standards of the good practices of hospital hygiene [22, 23, 24]. The five aiming aspects, likely to generate an environmental infectious risk during the CRF treatment by the extra-renal practice [20] are: establishment of the hemodialysis center, the buildings, the equipment, personnel, and technical conditions of operation.

This phase is presented in the form of synthesis grid, which states the criteria to be satisfied (referential) in each aimed aspect. Stay closer with the existing, emphasizing variations which are estimated on the basis of following criteria of evaluation: major gap (LMa); minor gap (LMi) and acceptable (Acc). For each slack, necessary measurements for an upgrade are taken. These measurements can be engineering changes, working instructions, procedures, recommendations, without affecting as much the aspects "and monitoring controls" in the strict sense.

It should be noted, that to minimize the risk of evaluation error, this stage is carried out by two groups of experts of the establishment, working distinctly on the basis of a single document, emphasizing after rapprochement, a common evaluation grid (table 1).

The analysis of the situation of the studied hemodialysis center, shows that the establishment under study, presents considerable deficiencies, primarily related: to an infrastructure not much adapted to this activity, no standardization of the practices, a total weakness of the provisions governing the institutions as well as the technical installations, the absence of a program of human resource valorization (sensitizing, formation, and motivation).

That is summarized primarily by not taking into account the human and organizational aspects. On the other hand the center is equipped with trademark machines "fresenius 4008B and 4008C", whose device of measurement, of

control are reliable as well as an installation of trademark water treatment " CAMRO " in good operating condition.

Table 1: Summary of the case study evaluation

Aspects	LMa		LMi		Acc	
	Criterion	%	Criterion	%	criterion	%
Location	0	0	0	0	2	100
Buildings	2	9	7	32	13	59
Equipement	1	7	4	29	9	64
Personnel	2	22	3	33	4	45
Operation	10	40	6	24	9	36
Total	15	22	20	29	35	50

3.2 Functional analysis.

The control of a system thus supposes that we have the required knowledge to suitably understand it and that we have defined and implemented the technical and organizational provisions allowing to control and supervise the operation and to react to any event.

Given that our system is regarded as sociotechnical complex [25] its modeling then requires the use of the systemic approaches which are mainly focused on the interaction of sub systems and more particularly on the flow analysis.

A Flow diagram (figure 2) it indicates the principal stages of the process used for the patient treatment by ERP. This treatment requires a chain of interventions with, at each stage, the risk of human error, and/or a material failure [26], can lead to infectious risks. Thus, the field of analysis defined by the working group goes from the patient arrival in the hemodialysis center to its exit, it includes the reception and the session preparation, the care of the patient carrying a transmissible germ, the preparation of the generator, the connection of dialyzed the dialysis progress, the end of meeting and disconnection of the dialyzed, the evacuation of the linen and waste, and finally the departure of the dialyzed.

For the examined hemodialysis process '12 'steps were obtained. Thus, each step can induce the risk of bacterial and / or viral infection that will be identified within a dysfunctional analysis. In addition, a mapping identifying all risky situations of transmission by the different actors (categories of personnel) to the patient is carried out effectively.

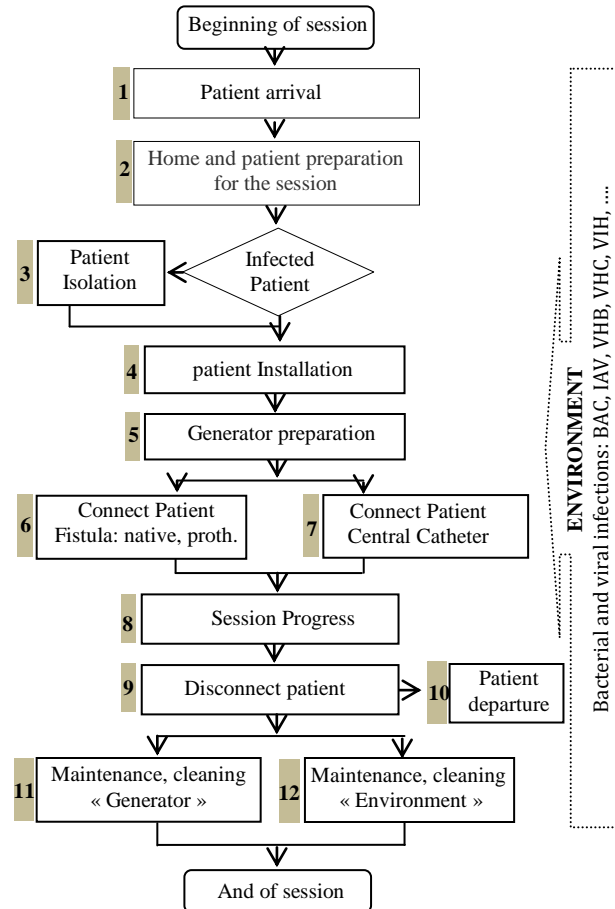


Fig.2. Flow diagram of 'Hemodialysis session'

3.3. Dysfunction analysis

3.3.1. Identification of the potential risks

The five determining elements in the control of the infectious risk (figure.3) for any hemodialysis activity are: the personnel (Labour), the air (Medium), the equipment (Material), and the products (Material) to which we add the organization and implementation methods.

So, the identification of the probable potential risks associated with the activity of care during all the chain of intervention for all its tasks rests on the situation analysis by taking into account all the intervening parameters.

For the activity of the hemodialysis center of the CHU Batna, 59 risky situations were identified and controlled; table.3 summarizes the main identified risks.

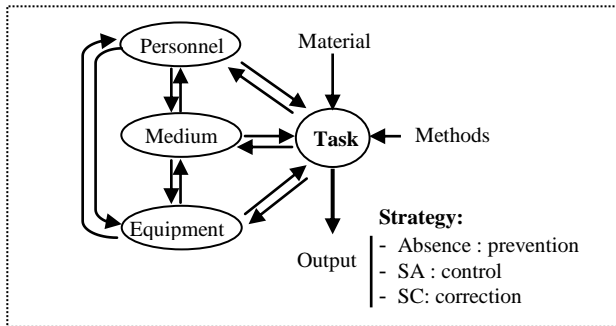


Fig 3. Infectious phenomena is occurred at the time of a care task

3.3.2. Estimate of the risk importance

It is an estimate carried out by a group of expert, qualifying the importance of the risks identified by ' Mi: Minor, Ma: Major, Cr: Critical (figure 4), in order to decide about the probable places where the critical points are situated, and about the required monitoring level to decrease the magnitude of the considered risk.

Probability of the event						
High	Ac	Mi	Ma	Cr		
Moderated	Ac	Mi	Ma	Ma		
Weak	Ac	Mi	Mi	Mi		
Negligible	Ac	Ac	Ac	Ac		
		Weak	Moderated	High	Severity of the consequences	

Fig 4. Estimation method of the danger importance [27]

Table 3. Principal identified risks

Identified risks	Summary of the principal identified risks
Risks of bacterial infections	<p>1. Infections related to vascular access:</p> <ul style="list-style-type: none"> The chronic hemodialysis generally implies a permanent vascular access generally an arteriovenous fistula, sometimes a catheter. These vascular accesses are exposed to risks related to their nature and their condition of use Contamination carried manually from environmental tanks (surfaces, objects.) or from infected patients, at the time of a hygiene rules breaking during the vascular access handling.

	<p>2. Infections not related to vascular access::</p> <ul style="list-style-type: none"> Dialysate contamination by its components (water, concentrated acid and bicarbonate). Contamination of the medicamentous solutions (nonspecific to the ERP practice) Contamination of the hydraulic system of the generator (bad disinfection..etc. Contamination system of the sewer drainage.
Risks of viral infections	<p>1. Contamination by virus VHB / VHC :</p> <ul style="list-style-type: none"> Internal contamination of dialysis generator by the blood of a previous dialyzed patient (flooding of the arterial or venous pressure sensor, dialysate hydraulic system). Injection of drugs or aqueous solutions that are contaminated by the patient blood carrying the virus from multi-purpose bottles. Virus presence on the vascular access or the site of injection of the extra corporal circuit of the receiving patient. This virus penetrates in the blood at the time of the vascular access puncture, of an injection or a time of taking. <p>Virus transmission by multiple ways: hands of the health care personnel, hands of infected patients, shared objects and medical devices, contamination of the generator surfaces and external components as well as the environment of the patient.</p> <p>2. Contamination by virus VIH.</p>
Risks of infections AEB	Not approached in this paper

3.3.3. Measure of control.

For each identified potential risk, measurements of control have been carried out in order to prevent or eliminate any risk for the patient to reduce it to an acceptable level. These measurements are referred primarily on the good operational practices and the good hygiene practices [28].

3.3.4. Determination of critical points.

A critical point must be at the same time controllable so that a specific action can be undertaken in order to reduce to an acceptable level and/or to eliminate the risk in this point and directed to take measurements in order to ensure itself of its controllability.

The question is to identify the risks which are not completely controlled by the hemodialysis center within

the framework of the good hygiene practices, and the good operational practices identified with a stage to which the application of a measurement for the control is essential. For that, the quotation method is used. To minimize the risk of error, the determination of the critical points by weighting, is carried out by two subgroups of experts, working distinctly on the basis of same criteria, emphasizing after rapprochement, a single result. It should be noted that the working group has exploited the existing literature, by taking into account the specificities and the proper conditions at the hemodialysis center of the CHU Batna.

The analysis made on 74 stages of the intervention chain in the hemodialysis center of the CHU Batna enabled us to determine 03 critical points to be controlled (table 4).

Table4. Critical control point.

Stage n°	Designation	CCP n°	Type
06	Connection of the patient.	CCP.1	Bacterial Viral
07	Disconnection of the patient.	CCP.2	
11	Disinfection of the hydraulic circuit.	CCP.3	

3.4. System of control.

The permanent control of the quality of the health care and in particular the patient safety and/or the health personnel pass by the implementation of a monitoring system of the hemodialysis activity in all its chain of intervention, thanks to observations or regular measuring. This system aims at being ensured of the implementation conditions of each identified preventive measure, by comparison to a target reference frame, and to set up pre-established corrective actions in the case of drift event. This system of control is exerted in controlled critical points, i.e. the tasks of the hemodialysis activity, which one must control in order to eliminate the risks or to reduce their probability of occurrence.

An example of control of the critical points is presented through the process "treatment of the hydraulic circuits of the generators" which requires the realization of three principal stages: cleaning, descaling, disinfection associated with a rinsing phase in the case of a chemical treatment. The stage n°11 'Disinfection' represents a critical point (CCP-3) corresponds to the probable risks of infections associated to this stage.

3.4.1. Critical and acceptable threshold

It is the separation of acceptable and the non-acceptable one by the establishment of a critical threshold for each critical point. If the monitoring reveals a slip towards the loss of control before exceeding the critical threshold, this threshold activate starts the preventive action and is called acceptable threshold (Figure 5).

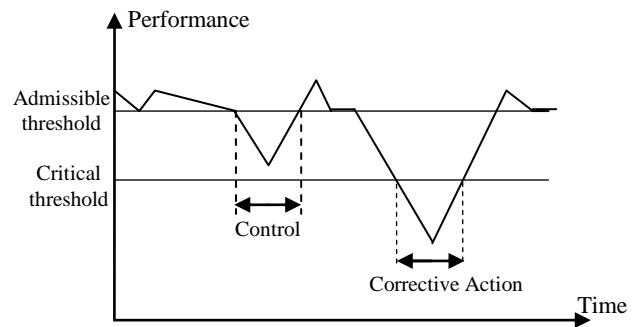


Fig 5. Critical limit and operational limit

An example of critical limit and operational limit (table 5) related to the CCP-3 'disinfection'.

Table 5. Critical threshold (CCP-3: disinfection).

Stage n°	CCP n°	Risk Description	Critical threshold
Disinfection of the hydraulic system of the generator	CCP-3	1. Temperature/time The non-respect of the scale (temperature / time) of disinfection can cause a bad treatment as well as the survival of pathogenic bacteria.	According to the type of the generators [29] According to Fresenius: - Heat: 84 °C – 15 mn to 30 mn. - chemical : 37 °C – 20 mn
		2. Concentration. The non-respect of the dose of the disinfection chemicals (chemical or thermochemical) can cause a bad treatment as well as the survival of bacteria.	According to the type of the products.

3.4.2. Monitoring system (preventive action).

The actions of monitoring allow the detection of the loss of control on the level of the CCP (table 6).

Table 6. Monitoring system (CCP-3: Disinfection).

<i>Monitoring Procedures «Preventive Action»</i> CCP-1 : Disinfection of the hydraulic system
<p>1. Control scale: temperature / time.</p> <p>Concerning the thermal disinfection, the effectiveness parameters i.e. the temperature reached in the circuit and the contact time, are measured and recorded throughout the treatment.</p> <ul style="list-style-type: none"> ▪ Monitoring 1 : The NGS must permanently supervise the real temperature displayed on the level of regulator PID. (Temperature regulation). Frequency: continuous. ▪ Monitoring 2 : The existence of a sound or visual alarm system starts in the case of anomaly of the disinfection cycle in particular the temperature and the pressure. Frequency: continuous. ▪ Monitoring 3 : The existence of a sound or visual alarm system and/or the stopping of the machine if the disinfection or the rinsing of the apparatus did not take place before a dialysis session. Frequency: continuous. ▪ Monitoring 4 : Once controlled, the dialysate passes in one to two ultra-filters. The role of these filters is the retention of bacteria or endotoxins, which can be developed within the generator, in spite of the disinfection procedures implementation [29]. <p>2. Control concentration.</p> <p>For chemical or thermochemical disinfection, the user must make sure at least, that the disinfecting quantity that allows obtaining the necessary concentration was indeed introduced into the circuits.</p> <ul style="list-style-type: none"> ▪ Monitoring 1 : For the generators equipped with automatic controller (flowmeter or level detector on the disinfecting internal vessel), the NGS must imperatively supervise these parameters. For the generators, which are not equipped with such means, the NGS must imperatively supervise by means of a visual reference mark located on the disinfecting container that the aspired volume corresponds to that required. Frequency: at each dosage operation. ▪ Monitoring 2 : The measurement of conductivity allows giving information on the concentration of the product present in hydraulic system.

<p>The NGS must supervise the measurement of conductivity permanently, if the system is equipped with a conduct meter. Failing that, samples will be taken for measurement. Any time the measurement of conductivity is not indicative, it reveals only one error of dilution. It is necessary to be ensured of the identity, the quality and the storage conditions of the chemicals.</p> <p>Frequency: continuous.</p> <ul style="list-style-type: none"> ▪ Monitoring 2 : The existence of a sound or visual alarm system starting at the moment of anomaly of the disinfection cycle. Frequency: continuous.
--

3.4.3. System of control of deviations.

Establishment of the procedures of deviation (corrective actions), for any deviation (loss of control), and for each critical point. The loss of control is regarded as a deviation compared to a critical limit for a CCP (table 7)

Table 7. System of control of deviations (CCP-3):

<i>procedures of deviation</i> <i>«corrective action »</i>
<p>1) In the case of deviation, not regulated by the automatic device, the NGS must :</p> <ul style="list-style-type: none"> a) Stop the process, and call upon the maintenance technician to carry corrections. b) Recording and reporting about the quality control and maintenance <p>2) In the case the recorded parameters show that the conditions for a given disinfection period are not met the NGS should:</p> <ul style="list-style-type: none"> a) Remake the disinfection process before any new activity of hemodialysis. b) Record: deviations and corrective actions. <p>- Responsible : NGS.</p>

3.4.4. Checking System.

It allows to confirm if the system of control is always valid and functions effectively (table 8).

Table 8. Checking System (CCP-3: disinfection).

<i>Monitoring Procedure</i> <i>« Checking »</i>
<p>1. Checking of the scale: temperature / time.</p> <ul style="list-style-type: none"> ▪ Checking 1 : Generally the generators are equipped with a system of measurement and traceability that allow recording the critical parameters of disinfection. The NGS must check that the required temperature, measured by the thermal probes, and the contact time was indeed reached during the cycle of disinfection.

Frequency: in the course and at the end of any operation of disinfection

2. Checking concentration.

▪ Checking 1 :

The NGS must check the evolution of the quantities of the consumed products per cycle. The number of cycles carried out with the same can or the order frequency of the products constitutes good indicators of operation.

Frequency: after each operation of disinfection.

3.4.5. System of traceability.

Establishment of a system (table 9) of documentation and recording of any action (preventive or corrective). Documentation and recording are essential, for the validity and the conformity of the implemented system.

Table 9. System of traceability (CCP-3: disinfection).

<i>Documentation and recording system</i>
- Report and recording of the deviations and corrective actions.
- Report and recording of the checks.

4. Conclusions

This item has no claim to be exhaustive; it is intended to be a work a more comprehensive approach to risk management. This is a field work and an experience sharing which aim at strengthening and consolidating of the current preventive approach to minimize this residual risk.

This approach usually involves an ongoing assessment of factors affecting the logistical and environmental characteristics that may affect infectious risk of the patient during the CRF treatment, as it involves checking and control for the sake of a continuous improvement. It must be constantly checked, supported and improved in order to offer a service quality of health care and safety.

The environment can play an important role to indirectly facilitate the transmission of microorganisms and to increase the risk of infections. In the hemodialysis center of Batna, space planning can be improved to ensure the quality of health care for improving patient safety: minimum separation between patients remains below the standard of a specific area, adjusting a specific area for drugs preparation and another one for the maintenance of reusable materials, sufficient water points for the equipment (one water point for four dialysis stations).

Standard precautions are far from being present in routine practice, especially for connection and disconnection of patients.

There is a need for the establishment of written protocols for the management of the dialyzed patient. There is a need to sensitize caregivers about the risk of cross-transmission.

As perspective, we can say that this approach requires a better consideration of the human factor, this by deploying the notion of variability of human performance, as is our current research work.

References

- [1] Hale, A. R. et Hovden, J. Chapter Management and culture : the third age of safety. A review of approaches to organizational aspects of safety, health and environment dans Occupational Injury. Risk Prevention and Intervention (Feyer, A. M. et Williamson, A., Éd.). Taylor & Francis, London, 1998.
- [2] Reason J. Managing the risk of organizational accidents. Aldershot: Ashgate; 1997.
- [3] Desroches A. Concepts et méthodes probabilistes de base de la sécurité. Paris: Lavoisier—Tec et Doc; 1995.
- [4] Rasmussen, J., & Svedung, I. Proactive risk management in a dynamic society. Karlstad-Suède: Swedish Rescue Services Agency, 2000.
- [5] Benzouai M, Mouss L.H, Smadi H. La maîtrise de la qualité dans l'industrie agroalimentaire, étude de cas. 4th International Conference on computer Integrated Manufacturing CIP'2007, Setif 03-04 novembre 2007.
- [6] Kassah-Louar A, Tobbi A. Séroprévalence des anticorps Anti-HCV, Anti-HIV et de l'antigène HBs chez les hémodialysés du CHU de Batna. JAM 2001, vol 4, 175-179.
- [7] Kassah-Louar A, Tobbi A, Benmehidi M. Séroprévalence des anticorps Anti-HCV, Anti-HIV et de l'antigène HBs chez les hémodialysés du CHU de Batna 1995-1997. RICAI 2001, 337/P2,214.
- [8] Kassah-Louar A, Tobbi A, Benmehidi M. Séroprévalence des anticorps Anti-HCV, Anti-HIV et de l'antigène HBs chez les hémodialysés du CHU de Batna (année 2003), Journées internationales de Khenchela. Mai 2007.
- [9] Kassah-Louar A, Tobbi A, Benbouza A, Millakhessou S, Chinar A. Hépatite virale «C» chez les hémodialysés chroniques du CHU de Batna, Prévalence et facteurs de risque (année 2009). La Revue médicopharmaceutique N°58 – 1^{er} trimestre 2011.
- [10] Circulaire DGS/DH/AFSSAPS No 2000–311 du 7 juin 2000 relative aux spécifications techniques et à la sécurité sanitaire de la pratique de l'hémodiafiltration et de l'hémodiafiltration en ligne dans les établissements de santé. Bulletin Officiel No 2000–25 www.sante.gouv.fr.
- [11] Canaud B. Hémodialyse : maîtrise des risques infectieux. In : Fabry ed. Maîtrise des infections nosocomiales de A à Z. Lyon : Health and Co; 2004: 374-384.
- [12] US RENAL DATA SYSTEM, USRDS 2003 Annual Data Report: Atlas of End-Stage Renal Disease in the United States, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2003. Disponible à: <http://www.usrds.org>, 2003.

- [13] Ayzac L, Beruard M, Girard R, et al. Dialin: Infection surveillance network for haemodialysis patients. First results. *Nephrol Ther* 2009, 5:41-51.
- [14] Hajjar J, Girard R, Marc JM, Ducruet L, Beruard M, Fadel B, et al. Intérêt de la surveillance des infections chez les hémodialysés chroniques en centre. *Bull Epidemiol Hosp* 2002; 3:10—2.
- [15] Hoen B, Paul-Dauphin A, Hestin D, Kessler M, EPIBACDIAL. A multicenter prospective study of risk factors for bacteremia in chronic hemodialysis patients. *J Am Soc Nephrol* 1998;9:869—76.
- [16] CCLIN Sud-Est. Site du réseau Dialin : <http://clin-sudest.chulyon.fr/reseaux/DIALIN/dialin.htm>. Saint-Genis-Laval; 2007.
- [17] Mayhall CG. Hospital epidemiology and infection control. Lippincott Williams & Wilkins, Philadelphia, London, 2004; 2060 p.
- [18] CCPS. – LOPA – Layer of Protection Analysis, simplified process risk assessment ,2001.
- [19] Décrets n° 2002-1197 du 23 septembre 2002 relatif à l'activité de traitement de l'insuffisance rénale chronique par la pratique de l'épuration extrarénale, 2002.
- [20] Décrets n° 2002-1198 du 23 septembre 2002 relatif aux conditions techniques de fonctionnement des établissements de santé qui exercent l'activité de traitement de l'insuffisance rénale chronique par la pratique de l'épuration extrarénale, 2002.
- [21] Arrêté du 25 avril 2005 relatif aux locaux, matériels techniques et dispositifs médicaux dans les établissements de santé exerçant l'activité « traitement de l'insuffisance rénale chronique par la pratique de l'épuration extrarénale, 2005.
- [22] Circulaire DGS/DH n° 98/249 du 20 avril 1998. Précautions générales d'hygiène ou précautions « standard » à respecter lors de soins à tout patient, 1998.
- [23] Comité technique national des infections nosocomiales : « 100 recommandations pour la surveillance et la prévention des infections nosocomiales ». Ministère de l'emploi et de la solidarité, secrétariat d'Etat à la santé et à l'action sociale, Paris, 121 pages, 1999.
- [24] Revue officielle de la société française d'hygiène hospitalière. Bonnes pratiques d'hygiène en hémodialyse, volume VIII, n 2, 2005.
- [25] Aloui. S. Contribution à la modélisation et l'analyse du risque dans une organisation de santé au moyen d'une approche système, Thèse de doctorat, Paris, 2007.
- [26] Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P. Acute renal failure – definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 2004;8:R204–12.
- [27] Division de l'alimentation et de la nutrition, systèmes de la qualité et de la sécurité sanitaire des aliments - manuel de formation, FAO, service de la qualité et des normes alimentaires, 1995, 6 : 148.
- [28] Bonnes pratiques d'hygiène en hémodialyse. Recommandations de la SFHH. Revue officielle de la Société Française d'Hygiène Hospitalière Volume XIII - N°2 - Avril 2005.
- [29] Merlo S, Donadey A, Coevoet B, Legallais C. Hemodialysis monitors: French market. *ITBM-RBM* 28, 2007 150–168.

Image enhancement using fusion by wavelet transform and laplacian pyramid

S.M.Mukane¹, Y.S.Ghodake² and P.S.Khandagle³

¹ Department of Electronics & Telecommunication
SVERI's, College of Engineering, Pandharpur

² Department of Electronics & Telecommunication
SVERI's, College of Engineering, Pandharpur

³ Department of Electronics & Telecommunication
SSPM's College of Engineering, Kankavali

Abstract

The idea of combining multiple image modalities to provide a single, enhanced image is well established different fusion methods have been proposed in literature. This paper is based on image fusion using laplacian pyramid and wavelet transform method. Images of same size are used for experimentation. Images used for the experimentation are standard images and averaging filter is used of equal weights in original images to blur. Performance of image fusion technique is measured by mean square error, normalized absolute error and peak signal to noise ratio. From the performance analysis it has been observed that MSE is decreased in case of both the methods where as PSNR increased, NAE decreased in case of laplacian pyramid where as constant for wavelet transform method.

Keyword: *image fusion, Laplacian pyramid, wavelet transform, Mean Square Error, Normalized Absolute Error and Peak Signal to Noise Ratio.*

1. Introduction

In image fusion is the process of combining relevant information from two or more images into a single image. The resulting image will be more enhanced than any of the input images. The concept of image fusion has been used in wide variety of applications like medicine, remote sensing, machine vision, automatic change detection, bio metrics etc. With the emergence of various image-capturing devices, it is not possible to obtain an image with all the information. Sometimes, a complete picture may not be always feasible since optical lenses of imaging sensor especially with long focal lengths, only have a limited depth of field. Image fusion helps to obtain an image with all the information. Image fusion is a concept of combining multiple images into composite Products, through which more information than that of individual input images can be revealed.

2. Motivation:

The concept of data fusion goes back to the 1950's and 1960's with the search for practical methods of merging images from various sensors to provide a composite image which could be better identify natural and manmade objects. In the past decade, medical imaging, night vision, military and civilian avionics, autonomous vehicle navigation, remote sensing, concealed weapons detection and various security and surveillance systems are only some of the applications that have benefited from such multi sensor arrays [3]. Motivation for image fusion is mainly the step toward recent technological advances in the fields of image fusion technique method. Improved robustness and increased resolution of modern imaging sensors and, more significantly, availability at a lower cost, have made the use of multiple sensors common in a range of imaging applications [4]. In this image fusion technique using laplacian pyramid and wavelet transform method increases the image quality of a fused image. GUI of these implemented method [10].

3. Image fusion techniques theory

These techniques both method having the two input images and these two images with laplacian pyramid and wavelet transform are used for fusion. From these two methods of image fusion we get a single fused image which has better quality and more enhanced image.

3.1 Image fusion by using dwt

The wavelet transform decomposes the image into low-high, high-low, high-high spatial frequency bands at different scales and the low-low band at the coarsest scale. The L-L band contains the average image information whereas the other bands contain directional

information due to spatial orientation. Higher absolute values of wavelet coefficients in the high bands correspond to salient features such as edges or lines [7]. First the images are transformed to the wavelet domain with the function `dwt2.m`; the `dwt2` command performs single-level two-dimensional wavelet decomposition with respect to either a particular ('wname','type') or particular wavelet decomposition filters (Lo-D and Hi-D) show in fig.1

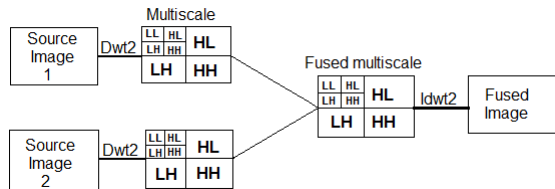


Figure.1 schematic of wavelet based fusion

The DWT is applied to both images and a decomposition of each original image is achieved. This is represented in the multi scale illustration where different bars (horizontal, vertical, diagonal and none) represent also different coefficients. The different boxes, associated to each decomposition level, are coefficients corresponding to the same image spatial representation in each original image, i.e. the same pixel or pixels positions in the original images. Only coefficients of the same level and representation are to be fused, so that the fused multi scale coefficients can be obtained [5]. This is displayed in the diagonal details where the curved arrows indicate that both coefficients are merged to obtain the new fused multi scale coefficients. This is applicable to the remainder coefficients. Once the fused multi scale is obtained, through the IDWT, the final fused image is achieved [6].

3.2 Image fusion by using Laplacian Pyramid

The image pyramid is a data structure designed to support efficient scaled convolution through reduced image representation. It consists of a sequence of copies of an original image in which both sample density and resolution are decreased in regular steps. These reduced resolution levels of the pyramid are themselves obtained through a highly efficient iterative algorithm. The bottom, or zero level of the pyramid, A_0 , is equal to the original image [2]. This is low pass- filtered and sub sampled by a factor of two to obtain the next pyramid level, A_1 . A_1 is then filtered in the same way and sub sampled to obtain A_2 . Further repetitions of the filter/subsample steps generate the remaining pyramid levels. To be precise, the levels of the pyramid are obtained iteratively as follows.

For $0 < l < N$:

$$A_l(i, j) = \sum_m \sum_n w(m, n) A_{l-1}(2i + m, 2j + n) \quad (1)$$

However, it is convenient to refer to this process as a standard DECREASE operation, and is given by $A_l = \text{DECREASE}[A_{l-1}]$. We call the weighting function $w(m, n)$ in equation 1 as the "generating kernel." For reasons of computational efficiency this should be small and separable. A five-tap filter was used to generate the Gaussian pyramid in Figure.2. Pyramid construction is equivalent to convolving the original image with a set of Gaussian-like weighting functions. The convolution acts as a low pass filter with the band limit reduced correspondingly by one octave with each level. Because of this resemblance to the Gaussian density function we refer to the pyramid of low pass images as the "Gaussian pyramid."

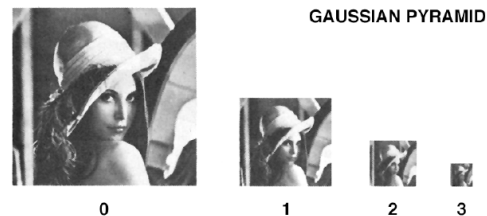


Fig.2 First six levels of the Gaussian pyramid for the "Lady" image. [11]

The original image, level 0, measures 256 by 256 pixels and each higher level array is roughly half the dimensions of its predecessor. Thus, level 5 measures just 8 by 8 pixels [11]. Band pass, rather than low pass, images are required for many purposes. These may be obtained by subtracting each Gaussian (low pass) pyramid level from the next lower level in the pyramid. Because these levels differ in their sample density it is necessary to interpolate new sample values between those in a given level before that level is subtracted from the next-lower level. Interpolation can be achieved by reversing the DECREASE process. We call this an INCREASE operation. Let $A_{l,k}$ be the image obtained by Increasing A_l k times.

Then

$$A_{l,k} = \text{INCREASE}[A_{l,k-1}]$$

Or, to be precise,

$$A_{l,0} = A_l, \text{ and for } k > 0,$$

$$A(l, k)(i, j) = 4 \sum_m \sum_n A(l, k-1)(2i + \frac{m}{2}, 2j + \frac{n}{2}) \quad (2)$$

Here in equation 2 only terms for which $(2i+m)/2$ and $(2j+n)/2$ are integers contribute to the sum. The INCREASE operation doubles the size of the image with each iteration, so that $A_{l,1}$ is the size of $A_{l,1}$ and $A_{l,1}$ is the same size as that of the original image. The levels of

the band pass pyramid, B_0, B_1, \dots, B_n may now be specified in terms of the low pass pyramid levels as follows:

$$B_i = A_i - \text{INCREASE}[A_{i+1}]$$

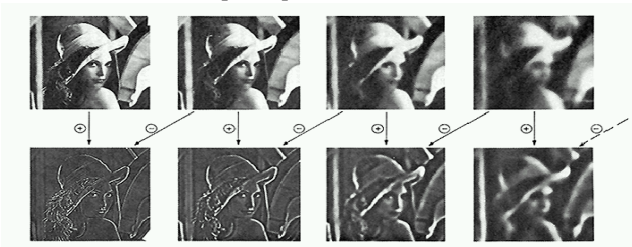


Figure.3 Generation of Laplacian images of 4 levels. [11]

The first four levels are shown in Figure.3 Just as the value of each node in the Gaussian pyramid could have been obtained directly by convolving a Gaussian like equivalent weighting function with the original image, each value of this band pass pyramid could be obtained by convolving a difference of two Gaussians with the original image. These functions closely resemble the laplacian operators commonly used in image processing for this reason we refer to the band pass pyramid as a "Laplacian pyramid." An important property of the laplacian pyramid is that it is a complete image representation: the steps used to construct the pyramid may be reversed to recover the original image exactly. The top pyramid level, B_n , is first expanded and added to B_{n-1} to form A_{n-1} then this array is expanded and added to B_{n-2} to recover A_{n-2} .

4. Experimental result and discussion

This experimental result and discussion depicts implementation for wavelet transforms and implementation of Laplacian pyramid

4.1 Implementation for wavelet transforms.

The discrete wavelets transform (DWT) allows the image decomposition in different kinds of coefficients preserving the image information [5]. Such coefficients coming from different images can be appropriately combined to obtain new coefficients, so That the information in the original images is collected appropriately Once the coefficients are merged, the final fused image is achieved through the inverse discrete wavelets transform (IDWT), where the information in the merged coefficients is also preserved [1]. The key step in image fusion based on wavelets is that of coefficients combination, namely, the process of merge the coefficients in an appropriate way in order to obtain the best quality in the fused image. This can be achieved by a set of strategies. The most simple is to take the average of the coefficients to be merged [6]. filter used in wavelet is db1

4.2 Implementation for laplacian pyramid.

Implemented laplacian pyramid fusion first construct the laplacian pyramid of two input images to be fused. Take the average of the two pyramids corresponding to each level and sum them. The resulting image is simple average of two low resolution images at each level. Decoding of an image is done by expanding, then summing all the levels of the fused pyramid which is obtained by simple averaging. A more efficient procedure is to expand B_n once and add it to B_{n-1} , then expand this image once and add it to B_{n-2} , and so on until level 0 is reached and g_0 is recovered. This procedure simply reverses the steps in Laplacian pyramid generation. The input arguments of this function are: Source images (im1, im2): must have the same size, and are supposed to be same size example both image of (m*n) matrix size. Number of scales (ns): an integer that defines the number of pyramid decomposition levels. The reduce operation is a two-dimensional convolution with the Gaussian filter followed by a down sampling by two. Expand, as the opposite operation of reduce, performs an up sampling by two followed by a two dimensional convolution with the same Gaussian filter [8]. Steps followed for fusion is given below table 1.

Table 1: steps for Laplacian pyramid fusion

$a11 = \text{reduce}(\text{image1}), a21 = \text{reduce}(\text{image2})$
$a11-1 = \text{expand}(a11), a21-1 = \text{expand}(a21)$
$B1 = \text{image1} - a11-1, B2 = \text{image2} - a21-1$
$B = \text{maximum}(B1, B2)$
Last level of the pyramid, $\text{fusion1} = \text{average}(a11, a21)$
Reconstruct fusion image, $\text{Fusion} = \text{expand}(\text{fusion1}) + B$

4.3 Performance analysis

Table 2 shows data base of input images and fused image. Image1 rice.png and image2 cameraman.tif are input images for fusion. Image3 and image4 fused images by wavelet transform, like wise image 5 and image6 fused images by laplacian pyramid method. Table 3 and 4 shows performance analysis. Performance analysis of the original images and fused image can be identify by three performance metrics Mean Square error, Normalized Absolute Error and Peak Signal To Noise Ratio shows in table 3 and 4

Table 2: Data base of input and fused images

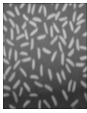





In put images Image1 (rice.png) 256*256 Image2 (Cameraman.tif) 256*256		
Fused images by wavelet transform.		
Fused images using laplacian pyramid.		

Table 3: performance metrics of cameraman .tif

Parameters	Results for cameraman.tif image		
	Image2	Fused image using wavelet	Fused image using Laplacian
MSE	851.17	645.5678	232.981
NAE	0.2501	0.2501	0.1310
PSNR	18.869	20.0654	24.4916

Table 4: performance metrics of rice.png

Parameters	Results for rice.png image		
	Image1	Fused image using wavelet	Fused image using Laplacian
MSE	1.0701e+03	684.165	283.9938
NAE	0.2498	0.2498	0.1298
PSNR	17.8706	19.8132	23.6317

5. Conclusion

Method implemented for laplacian pyramid and wavelet transform fusion is gives reliable fused image than the original image for finding the enhancement of the image taken parameter was mean square error, normalized absolute error of fused image reduce both kind of error and increase the peak signal to noise ratio. Image fusion is very much necessary in lot of applications to obtain the high quality image from two source images. There are different image fusions techniques are leading now days but in this seminar we have considered laplacian pyramid. Image fusion using laplacian pyramid and wavelet transform has been simulated using MATLAB software.

REFERENCES

- [1] Gonzalo Pajares, Jesus Manuel De Lacruz “A wavelet – based image fusion tutorial” the journal of the pattern recognition society, pattern recognition 37(2004)
- [2] Qiang Zhang, Long Wang, Huijuan Li Zhaokun Ma “ Similarity-based multi modality image fusion with shiftable complex directional pyramid” pattern recognition letters 32(2011)1544-1553
- [3] V. Aslantas , R.kurban “Fusion of multi- focus images using differential evolution algorithm” expert system application 37(2010)8861-8870.
- [4] Anjali Malviya, S.G.Bhirud “Image fusion of digital image” International journal of recent trends in engineering, vol.2, no.3, November 2009
- [5] Krishta Amoliny, Yun Zang, Peter Dare “Wavelet based image fusion techniques –an introduction review and comparison” department of geology and geometrics engineering.universityofnewBrunswick,sciencedirect.com(1 2 July 2002 available) .
- [6] Jing Tain, Lichen “Adaptive multi-focus image fusion a wavelet-based statistical sharpness measure” school of computer and technology, Wuhan University of science, china.sciencedirect.com (8 Jan 2012).
- [7] Shutao Li, Bin Yang “Multi-focus image fusion by combining curvelet and wavelet transform” pattern recognition letters 29(2008) 1295-1307
- [8] Se-hwan.yun, Jin Heon Kim and Suki Kim, Senior member ieee “Image enhancement using a fusion framework of histogram equalization and laplacian pyramid” contributed paper electronic version published 0098 3663/10/ 2010 IEEE
- [9] Zhigun Wang, Djemel Ziou, Costas Armenakis , Derenli and Aingaum Li “ A comparative analysis of image fusion methods” IEEE transaction on geosciences and remote sensing, vol 43 ,no 6, June 2005.
- [10] Daniel Muller, Anthomy Maeder and Peter O’shere “The graphical image fusion toolkit” GIFTMICCAI-2006 open science workshop, sciencedirect.com (June 2006).
- [11] Peter J. Burt and Edward H.Adelson” The Laplacian Pyramid as a Compact Image Code” IEEE Transactions on Communications, Vol. Com-31, No. 4, Pages 532-540, April 1983



Shailendra M. Mukane

Prof. Dr. Shailendra M. Mukane has completed B.E. Electronics Engineering (1993), M.Tech Power Electronics from VTU with Gold Medal (2001) and Ph.D. Electronics Engineering from Solapur University (2013). His total experience is 19 years and research areas include signal processing, image processing, and power electronics. He is fellow member of IEI Kolkata, Life member of ISTE and Associate Member of IEEE.



Yogesh S Ghodake

Yogesh S Ghodake has completed B.E Electronics and Telecommunication Engineering (2011), pursuing M.E Electronics and Telecommunication in SVERI's College of Engineering Pandharpur.

User Authentication with Adaptive Keystroke Dynamics

Shimaa I. Hassan¹, Mazen M. Selim², and Hala H. zayed³

¹ Department of computer systems, Faculty of engineering,
Benha university, Shoubra, Egypt

² Department of computer Science, Faculty of computers and informatics,
Benha university, Benha, Egypt

³ Department of computer Science, Faculty of computers and informatics,
Benha university, Benha, Egypt

Abstract

Recently, keystroke dynamics increasingly being a field of interest for researchers, where users can access different systems through their typing styles, which increases the level of security. This paper tends to implement a robust keystroke dynamics system; it tends to solve the problem of samples variations by using an adaptive threshold. The proposed system is evaluated using CMU dataset, and a new dataset created for this work. Results obtained are compared with others reported in literature and proved to have good performance.

Keywords: *Keystroke dynamics, Timing features, Distance based measures, Leave-One-Out-Method.*

1. Introduction

Biometric systems make use of the physiological and behavioral traits of individuals, for authentication purposes. Physiological traits as: fingerprints, voice, hand-geometry, face, iris, retina, palm-print ...etc., and behavioral traits as: gait, signature, keystroke dynamics, and voice [1].

Keystroke dynamics is the process of authenticating individuals based on their typing style. It is not what you type, but how you type [2, 3]. Recently, keystroke dynamics biometric systems have become the alternative of username/password scheme, which has many drawbacks: passwords may be forgotten, attacked, or shared, so the system will be in danger. A user's typing pattern may be unique because of similar neuro-physiological factors that make written signatures unique.

Unlike other biometric systems that usually require additional hardware and thus are

expensive to be implemented, biometrics based on keystroke dynamics are almost free i.e. the only hardware required is the keyboard [4, 5].

The problem of keystroke dynamics is that, it is a behavioral biometric; so there are large intra-class variations in person's typing patterns due to changes in emotional state, position of the user with respect to the keyboard, and type of keyboard used. The collected samples of persons need to be updated periodically [5].

The organization of this paper is as follows. Section 2 presents the types of features that can be extracted and also a brief of the previous work in keystroke dynamics. Section 3 presents the proposed system. Section 4 presents the experimental results, and Section 5 concludes the work.

2. Previous Work

The features extracted from keystroke dynamics pattern in most of researches are timing features. Fig. (1) shows the extracted timing features: [6]

1. Key Hold (KD): time between key pressed and key released.
2. Down-Down Key Latency (DDKL): time between two successive presses.
3. Up-Up Key Latency (UUKL): between two successive releases.
4. Up-Down Key Latency (UDKL): time between the current key release and the next key press.
5. Down-Up Key Latency (DUKL): time between the current key press and the next key release.

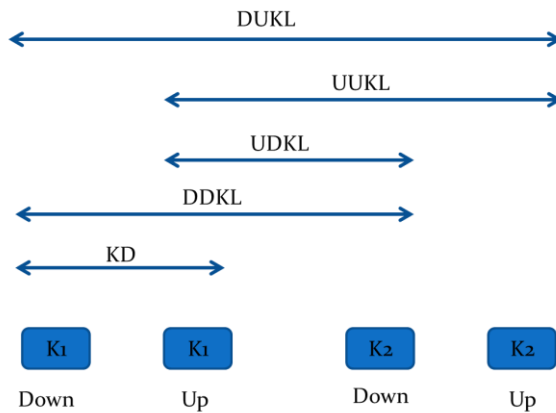


Fig. 1 Extracted features from keystroke timing patterns.

Gaines et al. (1980) were the first researchers showed that keystroke dynamics can be used for authentication [3, 2, 6]; they used long text (900-1200 words) and made their experiments using samples from seven users only.

Hosseinzadeh and Krishnan [6] used three keystroke features to authenticate users: KD, DDKL, and UUKL. UUKL is novel feature proposed during this work. These features were analyzed and modeled using GMM (Gaussian Mixture Modeling). The combination of the KD and UUKL features provided the best performance that led to an equal error rate (EER) of 4.4% based on a database of 41 users, each types 30 times.

Rybnik et al. [7] proposed a new approach to authenticate users using short fixed text. The extracted features were KD and UDKL. Classification of samples is based on k-nearest neighbor algorithm; the best accuracy obtained is 90.38% for 21 users.

Killourhy and Maxion [8] use a 14 keystroke dynamics detectors to authenticate users, 11 detectors was proposed by previous researchers, and 3 classic pattern recognition detectors (Euclidean, Manhattan, and Mahalanobis distance measures). Their data were collected from 51 individuals, each typed the same password 400 times along 8 sessions (50 times/session), 200 samples are used for training and the other 200 samples are used for test, the features extracted from each sample are: DDKL, UDKL, and KD. Scaled Manhattan provided the best results, and reduces EER to 9.6%.

Romain Giot et al. [9, 10] proposed a new method based on SVM (Support Vector Machine)

learning. Data were collected by allowing each user to type a fixed password "greyclaboratory". The features extracted from each sample were: UUKL, DDKL, UDKL, DUKL, and the total typing time. They use a population of 100 persons; each produced only 5 captures for the enrollment step. SVM with intelligent adaptation mechanism and the individual threshold produced the best results, which reduces EER to 6.95%.

Pin Shen Teh et al. [11] proposed a new system that uses two measures to calculate the similarity score between the two given samples: Direction Similarity Measure (DSM) [12] and Gaussian probability Density function (GPD). They evaluated their system over 100 persons; each typed 10 times their usernames, passwords, and a fixed phrase "the brown fox". The results were obtained by applying a two layer fusion approach on both GPD scores and DSN scores. The extracted features were KD, DDKL, UUKL, and UDKL. KD and UDKL yields the best results with all used fusion rules, using them with "And Voting Fusion Rule" produced EER near 1.4%.

Deian Stefan et al. [13] used keystroke dynamics for authentication and detecting imposters, they showed its robustness against forgery attacks. They presented a framework called TUBA for monitoring a user's typing patterns. They used the total typing time, KD, DDKL, UUKL, UDKL, and DUKL features to authenticate users. Support vector machine is used for classification. They evaluated their system using 20 users' keystrokes. The best result they obtained is 4.2% for average false positive rate.

Yu zhong et al. [14] evaluated a keystroke biometrics algorithms based on a new distance metric on the keystroke dynamics dataset created in (CMU Dataset) [8]. The new distance metric combined both Mahalanobis distance and Manhattan distance. Using the Nearest Neighbor classifier with the new distance metric achieved an average EER of 8.7%.

In previous work there are some constraints. In [7, 13], users authenticated through long phrases which is not real in case of passwords. Where in [6, 8, 14], users are asked to type several times in different sessions which could be refused by some users as it is time consuming and needs additional efforts. Also in [8, 9, 11], all users typed the same word, but in the real login system each user types his own password.

The main idea of our work is to allow users to access different systems by typing their own usernames and passwords as usual. Then, the users' typing styles features are extracted from their passwords, so there is no additional text required for authentication.

3. Proposed System

The block diagram of the proposed system is given in Fig. (2).

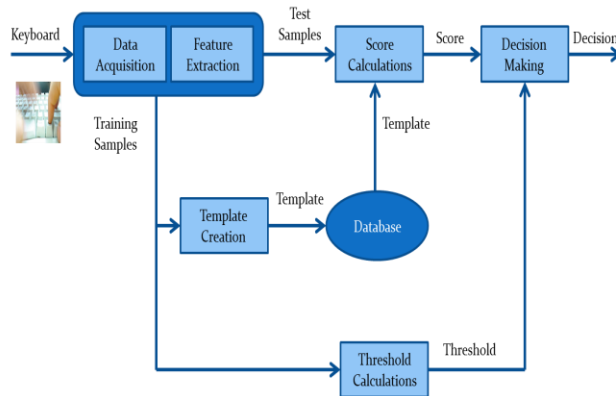


Fig. 2: Block diagram of the proposed keystroke dynamics system.

A new keystroke dynamics system was proposed. The following steps were used to implement the proposed system as follows:

1. The individual types their username and previously trained eight character password several times through separate sessions.
2. Features are extracted when individuals press and release keys.
3. User's Template and threshold was calculated through the extracted features.
4. Calculate the distance between template and the test samples to get the user's score
5. Finally, the user's score is compared against its threshold to make the decision.

3.1 Data acquisition and Feature Extraction

New dataset is created to evaluate the proposed system. A software application is implemented to acquire samples from individuals and extract their features, simply, user types his username

and password, and individuals can easily run this program on their own PCs or Laptops. Users are allowed to enter their own eight characters passwords containing only uppercase characters, lowercase characters, and numbers, where special characters are not allowed. Time stamps of each key press (Down) and release (Up) are stored in a log out file and used to calculate KD, DDKL, UUKL, UDKL, and DUKL.

3.2 Keystroke Dynamics Algorithms

Four distance based algorithms were used to evaluate the system: Manhattan, Manhattan with standard deviation, Euclidean, and Mahanabolis.

3.2.1 Manhattan Distance

The score is calculated as in Eq. (1) which represents Manhattan distance [15]:

$$M = \sum_i^n (x_i - y_i) \quad (1)$$

Where $x = (x_1, x_2, \dots, x_n)$ represents test vector and $y = (y_1, y_2, \dots, y_n)$ represents the mean vector of the training samples

3.2.2 Manhattan with Standard Deviation Distance (std)

The standard deviation of each feature is calculated as well [8]. Eq. (2) will be in the form:

$$Ms = \sum_i^n (x_i - y_i) / \alpha_i \quad (2)$$

3.2.3 Euclidean Distance

The score is calculated as the squared Euclidean distance between the test vector and the mean vector as in the following Eq. (3):

$$E = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (3)$$

3.2.4 Mahanabolis Distance

The standard deviation of each feature is calculated, where the Mahanabolis distance is presented by Eq. (4) [15]:

$$Mh = \sqrt{\sum_i^n ((x_i - y_i) / \alpha_i)^2} \quad (4)$$

3.3 Thresholds calculation

User's threshold is more recommended than global threshold in user authentication, it is proved to produce better performance. Leave-One-Out-Method (LOOM) [6] is used to calculate thresholds for individuals through the following steps:

1. Dividing the training space of (n) samples to one sample used as test sample, and (n-1) samples used to create the training sample.
2. Applying a distance measure (Euclidean for example) to calculate the distance between the selected test sample and the mean vector of the (n-1) training samples.
3. This process is repeated (n) times and produce (n) different thresholds for each feature vector.
4. The average of these (n) thresholds is calculated to produce the individual threshold.
5. These steps are repeated to calculate the individual thresholds for the other three distance measures.

3.4 Scores calculation

A match score is known as a genuine score if it results from matching of two samples of the same user, otherwise it is known as an impostor score if it involves comparing biometric samples of two different users [1].

In the proposed system two sets of genuine samples and one set of imposters are used to authenticate users, the scores are calculated through the following steps:

1. The individual types its username and eight character password.

2. Features are extracted from the typed sample to produce KD, DDKL, UUKL, UDKL, and DUKL feature vectors.
3. The new feature vectors are compared with those of the individual stored template using one the four distance measures.
4. The obtained scores for each feature are reported.
5. These steps are repeated to calculate the individual scores within the other three distance measures

3.5 Decision making

The proposed system is evaluated using: False Rejection Rate (FRR), which is the refused fraction of genuine individuals, and False Acceptance Rate (FAR), which is the accepted fraction of impostor individuals. Eq. (5) and (6) shows FRR and FAR respectively.

$$FRR = \frac{\text{Number of refused genuines}}{\text{Total number of genuines}} \quad (5)$$

$$FAR = \frac{\text{Number of accepted imposters}}{\text{Total number of imposters}} \quad (6)$$

The biometric system performance could be measured using Equal Error Rate (EER) which refers to the point on the ROC (Receiver Operating Characteristic) curve where the FAR and the FRR are equal [1, 3].

4. Experimental Results

Two datasets are used to evaluate the system; the first is CMU created by Kevin Killourhy [8], the second is created through this work. CMU dataset contains 51 individuals; each one typed the same password 400 times along 8 sessions (50 times/session). In this work, eight samples are used for each individual (one sample/session), six samples are used to create the training space, two samples are used to evaluate the system based on FRR, and two imposter samples are used to evaluate the system based on FAR. The extracted features for this dataset are: KD, DDKL, and UDKL and the previously stated matching algorithms were used to calculate scores and thresholds.

The most common approach for decision level fusion is majority voting (MV) [1]. If there are n features, the input sample is assigned an identity when at least k of the features agree on that identity, where $k = (n/2) + 1$ if n is even and $k = (n+1) / 2$ if n is odd.

Firstly, users are authenticated based on each feature separately; Table (1) shows the results of each feature, UDKL produces the best results is 8.8% for EER using Manhattan with standard deviation.

Table 1: EER of each feature for the CMU dataset

	Euclidean	Manhattan	Mahanabolis	Manhattan with std
KD	20.3	18.0	19.7	18.0
DDKL	17.2	14.7	15.3	12.3
UDKL	16.0	13.7	19.0	8.8

Then, individuals are authenticated based on different features combinations by comparing the average scores of combined features with their average thresholds. Finally, individuals are authenticated based on majority voting (MV), the best results is 7.0 % for EER using Manhattan with standard deviation and MV. Table (2) shows the best results for the features combinations.

Table 2: EER for CMU dataset based on two features and all features combinations

	Euclidean	Manhattan	Mahanabolis	Manhattan with std
KD & DDKL	16.0	14.5	10.3	8.1
KD & UDKL	14.9	13.0	10.3	8.9
DDKL & UDKL	16.7	14.3	16.7	11.3
KD & DDKL & UDKL	15.5	14.8	11.1	8.7
MV	15.5	12.1	12.5	7.0

EER is used to compare the results of the proposed system on CMU dataset with two existing systems: Kevin S. Killourhy (2009) [8] and Yu Zong (2012) [14], see table (3).

Table 3: shows the comparison between the proposed and other two systems based on EER

System	EER
Kevin S. Killourhy (2009) [8]	9.6
Yu Zong (2012) [14]	8.4
The proposed system	7.0

Fig. (3) shows a comparison among four distance measures using different features combinations based on EER.

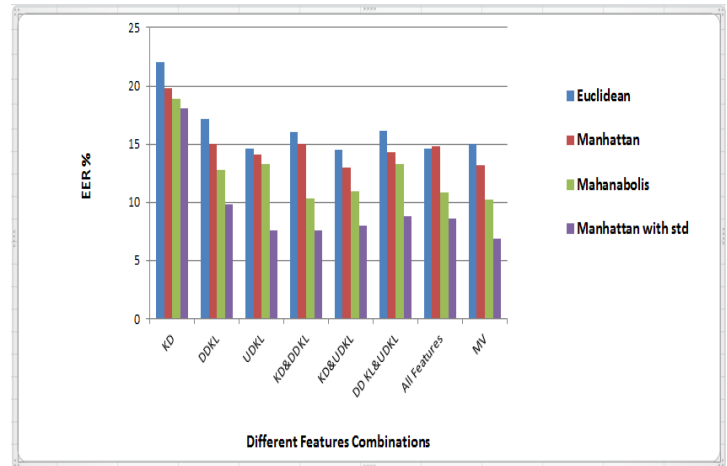


Fig. 3: a comparison among four distance measures using different features combinations based on EER for CMU dataset.

Fig. (4) shows the way to calculate the EER in the case of using Manhattan with standard deviation and MV on CMU dataset.

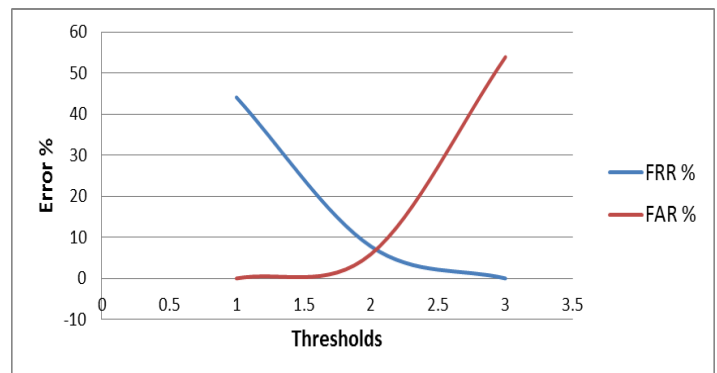


Fig. 4: ROC curve represents FRR and FAR for different thresholds using Manhattan with standard deviation and MV, the point of intersection represent the EER

New dataset of 62 individuals was created; each individual types his own eight character

password. Six samples are used to create the user template which stored in the database, and two genuine samples are used to evaluate the system, on the other hand two other imposter samples are used to evaluate the ability of the system to discover the forgery attacks. The features extracted are: KD, DDKL, UUKL, UDKL, and DUKL. The matching distance based algorithms are used to calculate scores and thresholds are: Euclidean, Manhattan, Manhattan with standard deviation, and Mahanabolis.

Firstly, users are authenticated based on each feature separately; Table (4) shows the results of each feature, UDKL produces the best results is 7.5% for EER using Manhattan with standard deviation.

Table 4: shows EER of each feature for the new dataset

	Euclidean	Manhattan	Mahanabolis	Manhattan with std
KD	16.5	15.0	15.0	14.7
DDKL	11.3	9.4	11.6	9.3
UUKL	12.0	9.5	11.3	8.3
UDKL	13.8	9.1	10.1	7.5
DUKL	10.1	9.0	8.2	7.8

Then, individuals are authenticated based on combinations of each two, three, and four features respectively, by comparing the average scores of combined features with their average thresholds. Table (5) shows the best results for the features combinations.

Finally, individuals are authenticated based on a combination of all features by comparing the average score all features with their average threshold, or by making vote (MV), the best results is 4.9% for EER using Manhattan with standard deviation and MV. Table (6) shows these results.

Table 5: shows EER for two, three, and four features combinations

	Euclidean	Manhattan	Mahanabolis	Manhattan with std
KD&UDKL	9.9	9.1	7.0	6.5
UDKL&DUKL	11.4	9.5	7.3	6.5
KD&DDKL&UDKL	9.3	8.0	6.7	5.7
KD&UUKL&UDKL	9.5	7.1	7.3	5.8
KD&DDKL&UUKL&UDKL	9.9	6.8	7.5	6.2
KD&DDKL&UUKL&DUKL	10.3	8.0	7.9	6.7

Table 6: shows EER of the new dataset based on MV and on all features combination

	Euclidean	Manhattan	Mahanabolis	Manhattan with std
MV	10.6	7.8	7.0	4.9
All Features	10.1	8.1	7.9	6.6

Fig. (5) shows a comparison among four distance measures using different features combinations based on EER for new dataset.

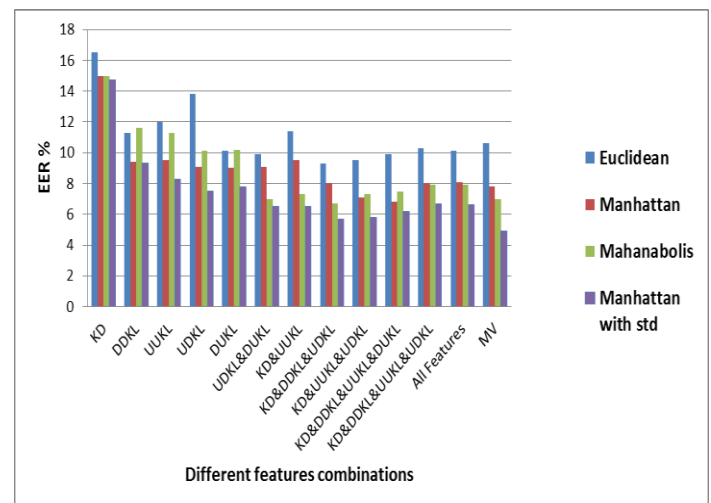


Fig. 5: a comparison among four distance measures using different features combinations based on EER for new dataset.

5. Conclusion

Keystroke Dynamics can be used as a digital signature to authenticate persons; it has no additional efforts from users to login systems, as users already type their passwords to be allowed to access. In some of the previous work users were authenticated by allowing them to type the same word or long phrases, while in the proposed system users are authenticated through their own eight character passwords; then timing features (KD, DDKL, UUKL, UDKL, and DUKL) are extracted for the typed characters. The system is evaluated based on each feature separately, and on different combinations of features.

The proposed system is evaluated using four distance measures: Manhattan, Manhattan with standard deviation, Euclidean, and Mahanabolis, for the matching process, taking the standard deviation into consideration increases the performance. Manhattan with standard deviation produces the best results as it takes into account the standard deviation of the training samples, so using it with the user based threshold calculated using LOOM could improve the system performance as it could solve the problem of large intra-class variations in user's samples. Two data sets were used in this work, the CMU data set and the other one is created via 51 individuals. Results obtained show a better performance while compared with the others.

6. References

- [1] A. K. Jain, P. Flynn and A. A. Ross, Handbook of Biometrics, Springer, 2008.
- [2] F. Monrose and A. D. Rubin, "Keystroke dynamics as a biometric for authentication," Science Publishers B. V. Amsterdam, The Netherlands, Feb, Elsevier, 2000.
- [3] M. Karnan and K. M. Akila, "Biometric personal authentication using keystroke dynamics: A review," Applied Soft Computing, Elsevier, 2011.
- [4] D. Jamil and M. N. A. Khan, "Keystroke Pattern Recognition Preventing Online Fraud," International Journal of Engineering Science and Technology (IJEST) vol. 3, March, 2011.
- [5] E. Lau, X. Liu, C. Xiao and a. X. Yu, "Enhanced User Authentication Through Keystroke Biometrics," Computer and Network Security Final Project Report, Massachusetts Institute of Technology, December 9, 2004.
- [6] D. Hosseinzadeh and S. Krishnan, "Gaussian Mixture Modeling of Keystroke Patterns for Biometric Applications," IEEE Systems, Man, and Cybernetics Society, November, 2008, Vol. 38, Toronto.
- [7] M. Rybnik, P. Panasiuk and K. Saeed, "User Authentication with Keystroke Dynamics using Fixed Text," International Conference on Biometrics and Kansei Engineering, 2009.
- [8] K. S. Killourhy and R. A. Maxion, "Comparing Anomaly-Detection Algorithms for Keystroke Dynamics," IEEE, Dependable Systems and Networks - DSN, 2009.
- [9] R. Giot, M. El-Abed and C. Rosenberger, "Keystroke Dynamics With Low Constraints SVM Based Passphrase Enrollment," IEEE International Conference on Biometrics: Theory, Applications and Systems, 2009.
- [10] R. Giot, M. El-Abed, B. Hemery and C. Rosenberger, "Unconstrained Keystroke Dynamics Authentication with Shared Secret," Computer & Security, ScienceDirect, 2011.
- [11] P. S. Teh, A. B. J. Teoh, T. S. Ong and C. Tee, "Keystroke dynamics in password authentication enhancement," Expert Systems with Applications 37 (2010) 8618–8627, Elsevier, 2010.
- [12] P. S. Teh, A. B. J. Teoh, T. S. Ong and H. F. Neo, "Statistical Fusion Approach on Keystroke Dynamics," Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, 2008.
- [13] D. Stefan and D. (. Yao, "Keystroke-Dynamics Authentication Against Synthetic Forgeries," Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2010 6th International Conference on.
- [14] y. Zhong, Y. Deng and A. K. Jain, "Keystroke Dynamics for User Authentication," Computer Vision and Pattern Recognition Workshop, IEEE

computer society conference, June, 2012.

- [15] S. Hocquet, J. Ramel, and H. Cardot."User Classification for Keystroke Dynamics," Seoul, Korea, Advances in Biometrics, International Conference, ICB, 2007.

Indiscernible Communication through ASCII Text Document/File (Communication in Veil)

Khan Farhan Rafat and Muhammad Sher

Department of Computer science, International Islamic University
Islamabad, 44000, Pakistan

Abstract

Internet has embarked comforting impact on people's lives across the globe as personal tasks such as cash transactions, mails, fax, data storage or retrieval etc. are now preferably being done electronically using personal computers/cell phones/TABs etc. This, however, has also elevated issues vis-à-vis availability, legitimacy, and renunciation of information causing discomfort in information security field because the number of cases including copyright infringement, heavy financial loss on account of e-theft of credit cards, impersonation through hacking of social IDs etc. are on increase, and calls for their immediate remedy. Steganography – the art and science of oblivious communication – has emerged as an exciting research area for expert and naïve academics now days to mitigate aforesaid security issues and is the prima facie of our research that focuses on evolving secure ASCII text-cover centric data hiding scheme for stealth communication.

Keywords: ASCII Text Steganography, Covert Channel, Eccentric way of writing, Stealth Communication.

1. Introduction

Take the situation of a person watching commercials on TV or adds in a magazine / NEWS paper where all of a sudden he/she gets pushed to buy that advertised item or get to that particular piece of interest without any compulsion or force or even without being asked to do so. This form of communication is referred to as oblivious communication and is induced in our lives to an extent that we take it as for granted and hardly realize its existence. However, when the intent is to hide personal data or some form of information from others, or to deliver it safely to another end without being noticed, the modus operandi is referred to as information hiding [1]. Steganography, a branch of information hiding, is not a new subject and dates back to 400 B.C. where Greeks pioneered the art [2] that has now been evolved into a science with digitization and the introduction of personal computers. Figure 1, derived from [3], illustrates on types of information hiding.

Earlier examples of steganography include wooden tablets, covered with wax, that carry hidden (engraved) messages and were retrieved by melting the wax [4]; an insurgent desirous to upheaval against the Persian king shaved the head of his

trusted slave, tattooed a secret message on it, waited for his hair to grow and sent him to his allied group of trusties in that territory where the slave's head was once again shaved to retrieve embossed message [5]. WORLD WAR - II saw use of invisible ink and micro dot techniques that highlights on the significance of use of Steganography [6].

Modern Steganography is more centered on exploiting human auditory-visual System (HAVS)'s limitations [7] where text, images and multimedia audio/video contents have become preferred means of stealth communication [8]. In fact Steganography emerged as an alibi in regions where Government / Corporate etc. has imposed ban on cryptography for public usage [9].

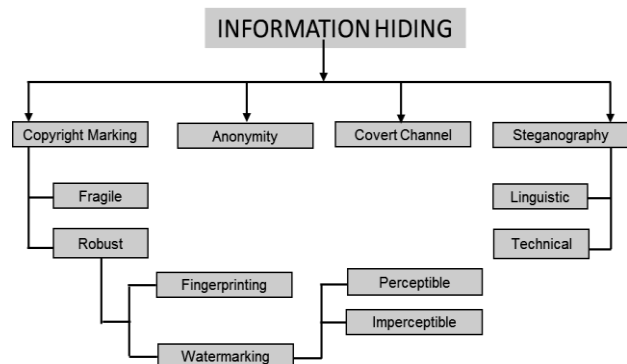


Fig. 1. Classifying Information Hiding.

ASCII Text-Cover centric steganography uses "text" as medium for covert communication. The distinguishing characteristic of text format from that of image, audio or video is that text files are viewed, saved and retrieved in a form analogous to what human eye perceives. Image, audio and video file formats, on the other end, have associated Meta data that dictates the manner in which that information is to be presented before end user. Further, changing a single bit of ASCII text character code results in another text code that has a dubious impact on the viewer e.g., misspelled word. It is because of the difficulty linked with ASCII text character codes that makes evolution of text – cover centered steganography a challenging task. Following briefly summarizes some text-cover based steganography schemes:

1.1 Text Content’s Manipulation

In [10] authors demonstrated variety of unconventional alternates for hiding secret information in ENGLISH text through deliberate modifications including syntax errors, use of acronyms, articulating document format etc. as shown below:

- Using intentional typographical mistake – writing “there” in place of “there”
- Preferring “yr” rather than “your” and “TC” in place of “Take Care”
- Formatting Text by inserting additional carriage returns or separating text into irregular paragraphs, or by adjusting line or word spacing.
- Through annotating text e.g., :) expresses pun
- Using multi lingual text – “we always commit the same mistakes again, and ‘je ne regrette rien’!”

1.2 Use of Blank/Space Character

Authors in [11] have suggested scheme that hides secret binary bit 1 using a single space while secret binary bit 0 depicts a double space. Following example illustrates the concept: paper must simply type your text into it.

Example. Let 00101110 be the secret message bits and let the text cover selected for bit hiding purpose be “A quick brown fox jumps over the lazy dog”. Going by analogy of suggested scheme secret binary message bit 0 will be replaced by double spaces in place of single white space character while no extra space is inserted for secret binary bit 1. After bit embedding process gets completed the resultant Stego Object takes the form as follows, where “ ” represents a double space:

A quick brown fox jumps over the lazy dog.

1.3 Paper Plot

This paper is planned as follows: Section 2 expounds on evaluation criteria while Section 3 elaborates on our proposed solution. Quantified test results and graphical outputs are illustrated in Section 4. Theoretical conception is given in Section 5. Section 6 highlights advantages and limitations of our proposed solution. Section 7 finalizes ongoing proceeding.

2. Evaluation Criteria

As stated earlier, steganography manipulates human’s HAVS limitations, hence the best evaluation parameter towards gauging any steganographic system after system’s security is its imperceptibility. However, it is apparent from [12] that Cachin’s [13] notion of perfect security (given below) may only exact on the similarity/difference

between cover text and stego object rather than describing system’s security:

$$D(P_c || P_s) \leftarrow \epsilon \leftarrow \sum_{q \in Q} P_c(q) \log_2 \frac{P_c(q)}{P_s(q)} \dots \dots (1)$$

, where P_c and P_s denotes probability distribution of Cover and Stego Object, It is obvious

$$\text{that } \frac{P_c(q)}{P_s(q)} = 0 \text{ for } \frac{0}{P_s(q)} \text{ and } \infty \frac{P_c(q)}{0}$$

Hence, to achieve information theoretic security we opted for the model proposed by [14] as illustrated in Figure 2. Quantified test results are tabulated using mean, variance and standard deviation. Additionally, to graphically illustrate bit embedding effects on text cover, probability distribution plots of cover text and stego object are contrasted using MiniTab16 [15].

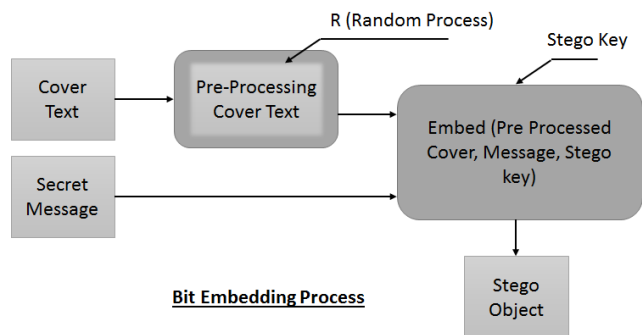


Fig. 2. Information Theoretically Secure Steganography Model.

3. Proposed Steganographic Scheme for ASCII Text Files

Instant inference on limitations of some existing methods for text steganography (Sect. 1 refers), less space manipulation scheme, include *perceptibility* (where erroneous words can catch viewer’s attention), *commencement of bit embedding from beginning of text that continues in the same direction till end of secret bits*, and *no indication regarding type (format) of data and its length that is embedded inside cover text*. Further, in the *absence of stego key*, a careful study of just few words/sentences may also reveal bit embedding algorithm. Based on above findings we wanted our proposed scheme to:

- Be in accordance with the evaluation criteria (Sect. 2 refers).
- Adhere to Kerchoff’s principle.

To meet our desired objectives we opted for the following (as explained in subsequent section):

- Use of symmetric stego key steganography.
- Reign in influence of stego key on bit embedding algorithm where 256-bit stego key length is preferred.
- Pre-processing of cover text before commencement of bit embedding process.

- To ensure 100% perceptibility we opted for insertion technique.

3.1 Design Considerations

With reference to peculiarity associated with ASCII text files i.e., contents are saved, and retrieved in manner in which these appear before human eye, we analysed ASCII text codes and found two such characters with codes 141 & 157 that even when made part of text file remains imperceptible (we shall refer to these ASCII codes as 'Stealth.Char') upon its retrieval. Visual Basic 6 used as tool for embedding, saving, retrieving and extracting *Stealth.Char* from ASCII text file where access mode for read/write operation was *binary*.

The constraint linked with *Stealth.Char(s)*, however, is that we can either interpret 141 as binary bit '0' and 157 as '1' or *vice versa* when hiding secret message bits which in absence of stego key can easily be comprehended by knowing bit embedding algorithm. Hence, to achieve variation in its interpretation during bit embedding we arranged *Stealth.Char(s)* as a 2 x 3 table and assigned these fixed values as shown in column two of the Table 1 leaving the third column as blank to be populated afresh with every new message for subsequent interpretation during bit embedding process.

Table 2: Quantified Test results

Stealth.Char	Assigned bit value	Key dependent interpretation
141	0	
157	1	

Next we added stego key byte values and reduced the result to modulo 32, the outcome of which was a pointer to the byte in stego key who's last bit i.e. 0/1 would decide either to retain or swap the prefixed bit pattern for interpreting *Stealth.Chars*. Figure 3 illustrates the said steps.

3.2 Processing the Stego Key

Keeping in view the salient characteristic of HASH function where a single bit change in input induces a change in at least half of the total output bits, the 32-bit stego key serves as input to SHA-2 HASH algorithm [16] and corresponding 256 bits stored/retained by counting number of ones' (binary bit 1). If *number of binary bit 1* \leq (*bits in secret message* + 64) then processes of bit embedding/extraction may commence. In case of otherwise, SHA-2 output serves as feedback (input) till the desired condition for binary bit 1 is met. For multiple iterations output from SHA-2 gets concatenated with previous generated HASH.

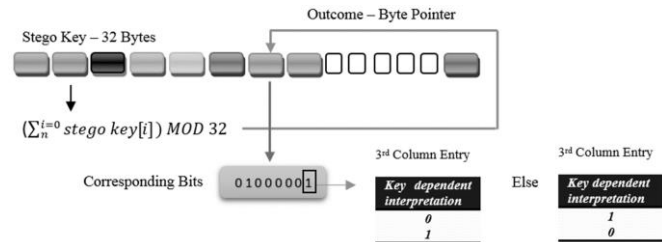


Fig. 3. Key dependent Stealth.Char interpretation in terms of binary bits.

3.3 Modus operandi

With preliminary work done, bit embedding and extraction processes are performed as under:

- Bit-Embedding:** Bit embedding commences by:
 - Encrypting secret message with stego key through XoR operation. The stego key is repeatedly used in case message length exceeds stego key length.
 - Translating secret message (*which is now encrypted*) into equivalent bits and storing its type (i.e. extension) and length in an eight byte header with four bytes reserved for each entry. The eight byte header is also translated into equivalent bits and attached as prefix to secret message bits.
 - SHA-2 generated HASH bits are traversed from left to right and the 'space' corresponding to binary bit 1 is marked for embedding encrypted bits.
 - In place of starting bit embedding from the beginning of ASCII cover text file, a random starting point for the said purpose is obtained using following equation:

$$\left(\sum_{i=0}^n \text{stego key} * (i+1) \right) \text{MOD } 65537) \text{MOD Total_Spaces} + 1 \dots (2)$$
 If the result is 1 the process of bit embedding continues normally otherwise it gets completed in two steps. In first step bit embedding commences from the point of insertion till end of encrypted message bits. In step two, bit embedding starts from the beginning of cover text file and continues till the point of commencement (Equation 2 refers) is reached or where the encrypted message bits get exhausted as the case may be.
 - Cover text is iterated taking encrypted bits in sequence. On finding the 'space' marked for bit embedding, *Stealth.Char* from column 3 of Table 1 corresponding to encrypted bit replaces the *Stealth.Char* attached with that 'space' during pre-processing stage. The process terminates when all encrypted bits gets embedded in cover text.
- Bit-Extraction:** Bit extraction commences by processing stego key (as in Sec. 3.2) followed by:
 - SHA-2 generated HASH bits are traversed from left to right and the 'spaces' in stego object corresponding to

binary bit 1 are marked as location containing hidden (encrypted) bit.

- b. Point for traversing stego object for extracting hidden bits is obtained as given in para 1 (d) above.
- c. The Stealth.Char(s) attached with each of the marked 'spaces' are replaced by the corresponding bits given in column 3 of Table 1.
- d. First 32 bits thus extracted gives hidden message type while the next 32 its length.
- e. Remaining extracted bits are XoR-ed with the stego key to get decrypted output.
- f. The output (hidden message) is then translated into equivalent bytes and saved in appropriate file format.

4. Test Results

In one of the experiments the perceptibility of stego object contrasted with cover (extracted from <http://en.wikipedia.org/wiki/Steganography>) and pre-processed cover text is illustrated vide Figures 4 – 6, Jaro-Winkler distance [17][18] computed equated to **0.9563** while similarity/difference observed between the three is shown by plotting probability distribution graphs using MiniTab 16 for which their mean, variance and standard deviation were calculated as shown in Table 2 and illustrated in Figures 8 - 10 respectively. We experimented with 75 ASCII cover text files of varied lengths and found the results close to those as exemplified.

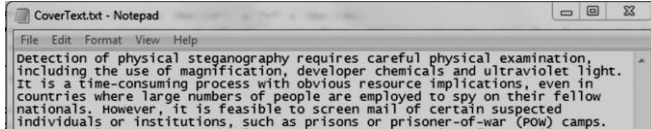


Fig. 4. ASCII Cover Text.

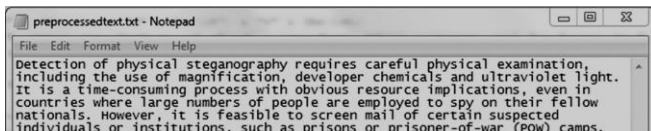


Fig. 5. Pre-Processed ASCII Cover Text.

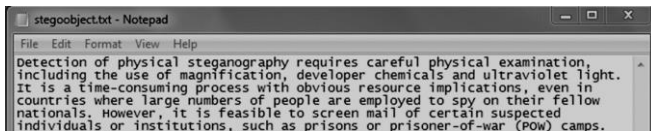


Fig. 6. Stego Object.

Figure 7 shows directory listing for the aforementioned three files where each is shown to have a file length of 3KB:

Local Disk (D:)	CoverText.txt	5/30/2013 12:56 PM	Text Document	3 KB
New Volume (E:)	preprocessed.txt	5/30/2013 1:23 PM	Text Document	3 KB
Local Disk (F:)	stegoobject.txt	5/30/2013 1:35 PM	Text Document	3 KB

Fig. 7. Directory Listing for Cover, Pre-Processed Text and Stego Object.

Table 2: Quantified Test Results

COMPUTATION	COVER TEXT	PRE-PROCESSED COVER TEXT	STEGO OBJECT
Mean	0.092388	0.099824	0.099765
Variance	0.881894	1.142325	1.136536
STD	0.029696	0.033798	0.033712

5. Theoretical Conception

For security to prevail the uncertainty about bit embedding must not get revealed merely on the basis of knowledge about cover text and stego object. Undoubtedly if 'Eve/Wendy' can spot differences between cover text and stego object then they can also break the system. This, however, can only be possible when the differences are caused by bit-embedding alone. Hence, security in steganography can be achieved through arbitrary selection of cover text and then pre-processing it via some hard to predict random process before applying bit-embedding over it under the control of stego key but without compromising on its perceptibility.

6. Advantages and Limitations

1. Advantages of our proposed scheme include:
 - i. Imperceptibility of stego object.
 - ii. Information Theoretic Security.
 - iii. Hidden message's type and length known before commencement of bit-extraction process.
 - iv. Key dependent arbitrary starting point to initiate bit embedding.
2. Some of the limitations are:
 - i. Increased stego object file size *equivalent to number of spaces in cover file*.
 - ii. Opening the stego object via applications other than Windows 'Notepad' may result in unintelligible text.

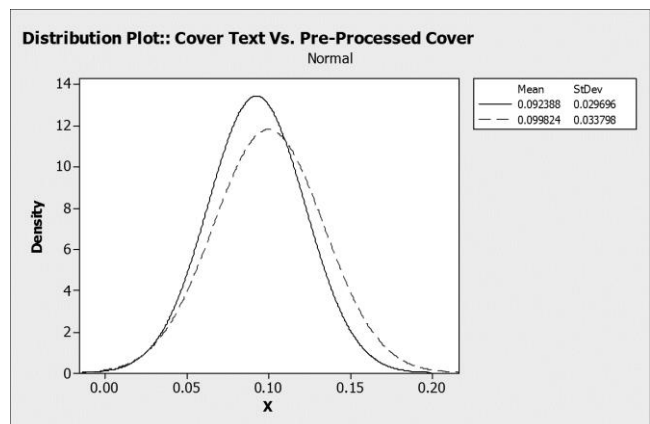


Fig. 8. Contrasting Probability Distribution Plots of Cover Text and Pre-Processed Cover Text.

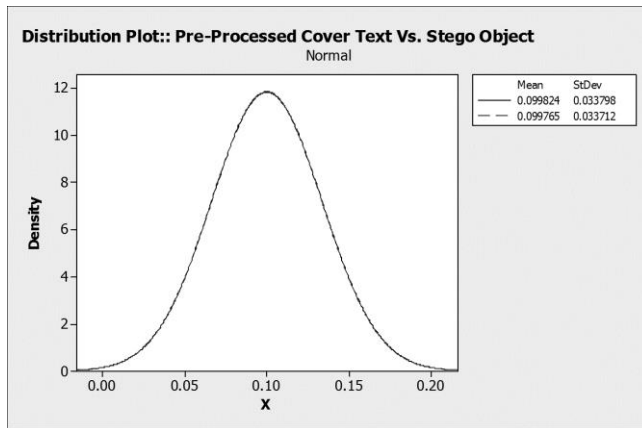


Fig. 9. Contrasting Probability Distribution Plots of Pre-Processed Cover Text and Stego Object.

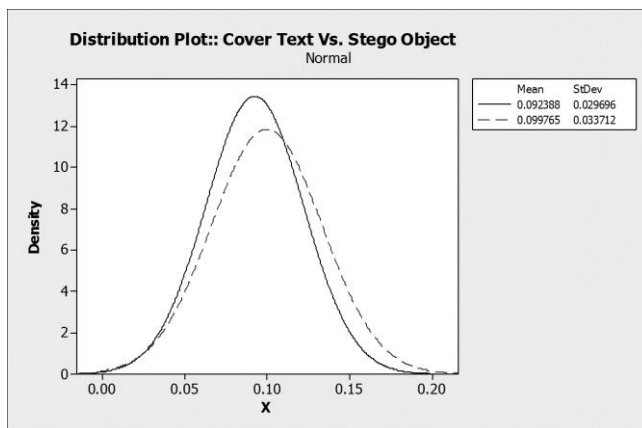


Fig. 10. Contrasting Probability Distribution Plots of Cover Text and Stego Object.

7. Conclusion

Less printing cost, high transmission efficacy, low resource occupancy and completeness in its semantics are some salient attributes that has made ASCII text documents the most commonly opted type of media in communication. However, lack of redundant information as well as non-alterability of alphabets while writing a character or word has made ASCII text document a difficult but challenging choice as Carrier/Cover for the purpose of information hiding in contrast to other media such as video, audio, image etc. having meta data that is/can be easily manipulated for the said purpose.

This research presented a secure steganographic scheme by inserting indiscernible characters corresponding to secret message bits in ASCII text document/file and is also in accordance with Kerchoff's principle.

References

[1] Pfitzmann, B., "Information Hiding Terminology," Proc. of First Internet Workshop on Information Hiding, pp. 347-350, Cambridge, UK, 1996.

[2] Stefan Katzenbeisser and Fabien A.P. Petitcolas, Introduction to information hiding. In Information Hiding: Techniques for Steganography and Digital Watermarking, Artech House. 1-14, Boston: 2000.

[3] Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, Information Hiding - A Survey, Proceedings of the IEEE, special issue on protection of multimedia content, 87(7):1062-1078, July 1999.

[4] Herodotus, The Histories, Penguin Books, London, 1996. Translated by Aubrey de Selincout

[5] N. Provos and P. Honeyman, "Hide and Seek: An Introduction to Steganography", IEEE Security & Privacy Magazine, Vol. 1, issue 3, pp. 32-44, June. 2003.

[6] N.F. Johnson and S. Jajodia, "Exploring Steganography: Seeing the Unseen," Computer, vol. 31, no. 2, pp.26-34, 1998.

[7] Der-Chyuan Lou and Chia-Hung Sung. "A steganographic scheme for secure communications based on the chaos and Euler Theorem", IEEE Transactions on Multimedia, Volume 6 Issue 3, pp. 501-509, June 2004.

[8] Miroslav Goljan and Andreas Westfeld. "Secure Steganography in Multimedia Content", EURASIP Journal on Information Security, 2009:257131 doi:10.1155/2009/257131.

[9] Jenny Shearer and Peter Gutmann. Government, "Cryptography, and the Right to Privacy." Journal of Universal Computer Science (J.UCS), Volume 2, No.3, p.113, March 1996.

[10] Mercan Topkara, Umut Topkara, Mikhail J. Atallah. "Information Hiding through Errors: A Confusing Approach", 2007. Internet: http://umut.topkara.org/papers/ToToAt_SPIE07.pdf, [July 12, 2012]

[11] Bender, W., Gruhl, D., Morimoto, N. & Lu, A. "Techniques for data hiding", IBM Systems Journal, Vol 35, pp. 313-336, 1996.

[12] Khan Farhan Rafat and M.Sher, "On the Limits of Perfect Security for Steganographic System", International Journal of Computer Science Issues, Vol. 4, July, 2013. www.ijcsi.org. **Accepted – Under Publication**

[13] Cachin, C. "An Information-Theoretic Model for Steganography," Proceedings of the Second International Workshop on Information Hiding, vol. 1525 pp. 306-318, Lecture Notes in Computer Science, Springer, 1998.

[14] J.Z Llner, H. Federrath, H. Klimant, A. Pfitzmann, R. Piotraschke, A. Westfeld, G.Wicke, G.Wolf. "Modeling the security of steganographic systems," Proc. 2nd Workshop on Information Hiding, pp. 345-355, LNCS 1525, Springer-Verlag, Portland, 1998.

[15] Minitab 16. <http://www.facebook.com/Minitab> [June 6, 2012]

[16] T Hansen - 2006, "Secure Hash Algorithms (SHA and HMAC-SHA)," <http://tools.ietf.org/html/rfc4634> [june 2012]

[17] Jaro, M. A., "Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida," Journal of the American Statistical Association 84:414-420, 1989.

- [18] Winkler, W. E., "The state of record linkage and current research problems," Statistics of Income Division, Internal Revenue Service Publication R99/04, 1999. <http://www.census.gov/srd/www/byname.html>.



KHAN FARHAN RAFAT is a Ph.D. Scholar at International Islamic University, Islamabad – Pakistan. He did his MCS from Gomal University, D.I.K. followed by MS in Telecommunication Engineering from UMT, Lahore – Pakistan. A veteran of information security having an experience of almost about 24 years has worked in varied roles in areas that are not limited only to programming, evaluation & analysis of Software/Hardware based security modules, and formulating security policies.



Professor Dr. Muhammad Sher is Dean Faculty of Basic and Applied Sciences at International Islamic University, Islamabad – Pakistan. He received B.Sc. degree from Islamia University Bahawalpur and M.Sc. degree from Quaid-e-Azam University, Islamabad, Pakistan. He did Ph.D. in Computer Science and Electrical Engineering from TU Berlin, Germany His area of research is Next Generation Networks Security. An eminent Scholar who has a number of research publications to his credit.

BIOMETRIC SYSTEM BY POINTING FINGER

Eugène MBUYI MUKENDI¹, Jean Didier BATUBENGA MWAMBA²

¹ Professor University of Kinshasa, Computer Science Department, DR Congo

² University of Kinshasa, Computer Science Department, DR Congo

Kinshasa Computer Science Laboratory.

Abstract

Biometrics is an important application of pattern recognition, the cross-domain path between data analysis and artificial intelligence. Technological change and the growing computerization of society have paved the way for biometric systems. They rest on sequence procedures that always follow the same development: biometric data is converted to digital data and stored in memory serve as biometric reference of the individual concerned. A biometric system helps to recognize a person whose identity has been previously stored in a database. In this paper, we present an automatic pointing system which allows identifying personnel from his fingerprint to know the time he/she attaches to the work. It is an authentication work aiming at verifying the identity of personnel during pointing in a company.

Fingerprint recognition's algorithm has been suggested by A.K. Jain who is probably the most known.

Keywords: biometrics, pattern recognition, identification, authentication.

INTRODUCTION [ABD1], [BAT2], [BOU3], [CHA4]

1.1. Biometrics

Biometrics is a branch of pattern recognition which always generates more interest. Biometric systems help to automatically identify people using physical or behavioral characteristics such as face, fingerprint, signature or walking. In other words, biometric refers to the identification of an individual based on morphological, biological or behavioral characteristics. These biometric characteristics are unique and specific to an individual and there is little possibility that other people may replace them. It is then more powerful in terms of safety.

Biometric data analysis can be classified into three broad categories:

- Analysis based on morphology: fingerprint, hand geometry, facial features, vein of the retina, iris, etc...
- Analysis of biological evidence: saliva, smell, urine, blood, DNA, etc...

- Analysis based on behavior: signature tracing dynamic, hits on a computer keyboard, voice, etc ...

These biometric features cannot be easily stolen, falsified or shared.

They must be reliable and secure for the recognition of people than traditional methods. Therefore, they must meet certain criteria for a broad reliability of biometric systems. They must be:

- Universal: must exist in every person
- Unique: to help distinguish one individual from another
- Permanent: stability over time
- Measurable: sensor technology exists that allows a comparison to the days to come
- Recordable
- Unfalsifiable.

1.2. Biometric system

1.2.1. Operating modes of a biometric system

Overall, a biometric system consists of two phases: an enrollment phase and a recognition phase.

Enrollment

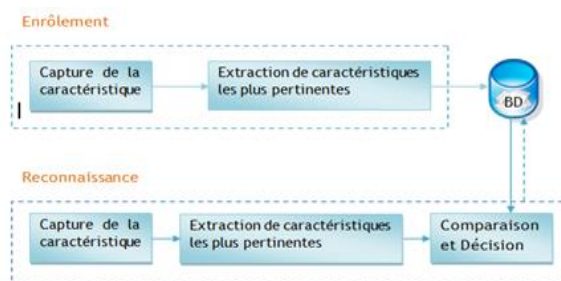


Figure 1. Architecture a biometric system

✓ Enrollment

This phase consists of creating a biometric template of an individual who will serve as a reference for recognition. For this purpose, the biometric characteristic of the individual is captured

by a biometric sensor, and then represented in digital form and then stored in a database. Ordinarily, we do not work directly on these raw data because they contain unnecessary information to recognition. Indeed, from this data, we do extract just the "relevant" parameters, this helps to significantly reduce the size of backup data.

✓ Recognition

During the recognition phase, the biometric characteristic is captured and the relevant parameters are extracted as in the enrollment phase. The follow-up of the recognition depends on the operating mode of the system. If you are in identification mode, the system will compare the captured signal with all templates contained in the database. Then it will pull the nearest signal template to answer question such as: "Who am I?" It is a very difficult task because the database may contain thousands of individual information. You lose a lot of time calculating all the possible comparisons. On the other hand, in verification mode (authentication), the system will compare the signal with one of the templates in the database with the aim of answering the question: "Am I the person I pretend to be?"

1.2.2. Evaluation of performance of a biometric system

Unfortunately, biometrics has a major drawback. In fact, none of the measures used reveals to be totally accurate because it is indeed a major feature of any living organism: it adapts to the environment, grows old, undergoes more or less important trauma, and therefore the measurements change. But it must nevertheless do well to gain recognition, and in reality it will work in most cases because the system allows a margin of error between the measurement and reference. Indeed, manufacturers, developers do not seek absolute security, they want something that works in practice. Thus we find the definition of the reliability of a biometric recognition system which is characterized by two measures of error. These two measures are the False Rejection Rate (FRR) and False Acceptance Rate (FAR). Manufacturers, developers of biometric recognition systems are seeking to reduce the FRR, while maintaining a relatively low rate of FAR. Therefore a compromise needs to found between FRR and FAR. In fact, these two measures of error are related, and there exists an equilibrium point between both of them for each biometric system. The graph, in which the two curves called ROC (Receiver Operating Characteristic) are shown, is purely demonstrative. Indeed, each measurement is based on the decision threshold of the system, a value determined by the designer or the system operator, who defines point when a match is made. Delta (Δ) represents the margin of error allowed by the system, varying from 0 to infinity. The more the allowable error margin is important, the more the false acceptance rate increases, that is to say we will accept more and more people who are not authorized (and therefore the safety of the system decreases).

However, if the rejection rate of authorized persons also decreases, this makes the system more functional and better meets user expectations. At the other extremity, if the error margin accepted by the process of biometric measurement is decreased the rate of the two trends are reversed: we will less and less accept people trying to cheat but we will also, at the same occasion, have a rejection rate on authorized persons which will be too large to be tolerated in most cases. The usual compromise is to consider the junction of the curves, that is to say the x point where FAR and FRR are equal and at the same time to a minimum. The more x value is low, the more the system is reliable, because there is a good equilibrium of sensitivity. Besides these two error rate, FAR and FRR, the failure rate to enrollment is equally used to determine the accuracy of a biometric system.

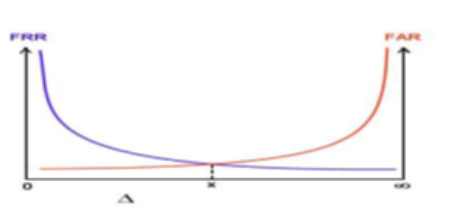


Figure 2. ROC Curve (Receiver Operating Characteristic)

- The Rate of False Rejection (FRR): it expresses the percentage of legitimate users rejected.
- The False Acceptance Rate (FAR): it expresses the percentage of impostors accepted.

$$FRR = \frac{\text{false rejections number}}{\text{clients number}}$$

$$FAR = \frac{\text{false acceptations number}}{\text{impostors number}}$$

1.2.3. Technique based on fingerprint

This is the oldest and most widely used identification technology. It represents more than forty-five percent of applications. The basic data in the case of the fingerprint is design represented by the ridges and valleys of the skin. This design is unique and different for each individual. In practice, the extraction of biometric data using this technique involves the imposition of a finger on a fingerprint reader.



Figure 3 biometric fingerprint reader

I.3. Fingerprints Recognition

The first traces of use of fingerprints were discovered in Egypt and date back to the time of the pyramids, more than 4,000 years ago. The Chinese have also used this medium very early to sign official documents, but they probably do not know that fingerprints were unique to each person and allow reliable identification. In 1856 the Englishman **William Hershel**, after using the fingerprint as a signature on Indian population he headed, began to realize that fingerprints were unique and constant over time.

In 1888, the British **Francis Galton** published a study on fingerprints where he established their characteristics (the notions of uniqueness, invariance, minutiae, classification). In 1901 the technique of identification through fingerprints was officially adopted in England in the judicial system. This technique was then greatly expanded in criminal investigations.

II. IMPLEMENTATION OF THE SYSTEM [CHA5], [CHA6], [MBU10]

❖ Identification

Each person must be characterized by biographical data (last name, middle name, first name, date of birth, sex, address, nationality, etc...) and biometric data (fingerprint).

❖ Fingerprint Capture (Enrollment)

The biometric data is captured directly from a fingerprint reader. The system should only allow saving the information necessary for the recognition and not to make the data base heavy.

❖ Search

Various registered persons can be identified from the data managed by the system. No matter what data it is biographical or biometric it can provide research services. We then speak of matching when it comes to biometric data.

❖ Payment Processing

The system must be able to calculate wages and provide employees payroll.

❖ Sending SMS

The system must be capable of processing payment. By the moment all calculations were well made, the system must be capable from a click to send SMS to the relevant staff to inform them of their situation.

❖ Statistics' Treatment

The system must be able to produce statistics on all employees. The number of days when one actually worked per month, per year, etc...

II.1. Collection of operational needs

❖ Security

It is the system administrator which manages the access rights. When connected, each user must be recognized from the system

and by his/her fingerprint and access only information that is necessary for him depending on the rights that they have.

II.2. Description of the system context

Now that the collection of needs is carried out, description of the context of the system can begin. It consists of three successive operations:

- Identification of actor,
- Messages identification,
- The realization of context diagrams.

II.3 Identification of actors

An **actor** is a user who always has the same behavior vis-à-vis a use case. This is an external component (hardware or other system device) which interacts with the system or a role that a user plays in relation to the system. A single user can play multiple roles as multiple users can play the same role. For example, the administrator of a messaging system can also be a user of the same messaging. He will be considered, as an actor of the system, in the administrator role on the one hand and that of user on the other hand.

Considering this definition, we identify the following actors:

Actors	Role
Administrator	Manages access to all configuration settings of the application. He handles the recruitment of employees. It processes the payment and prepares payroll. He consults statistics on all staff.
User	Role played by all those who have access to the application, but not accessing all the features of the application.
Fingerprint reader	It captures digital fingerprints
GSM Modem	Send SMS to personnel

Table 1: Identification of actors

II.4 Functional needs Capture

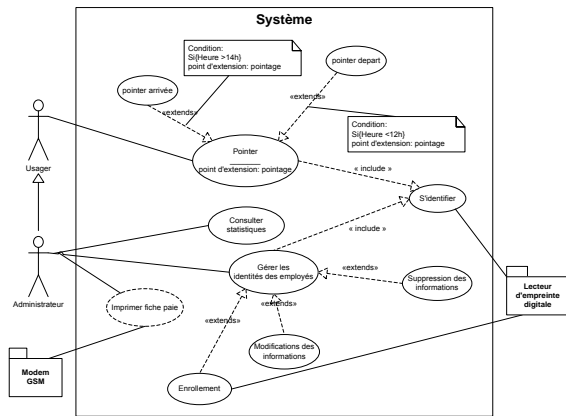


Figure.4 diagram of use cases

II.5. Application Presentation

To avoid overloading the paper, we present a number of screen captures of our application. A simple click on the executable of the application launches a splash screen. For security reasons when the application is provided for the first time, it contains no information. Thus we have provided a form of authentication where the user enters the password and login.



Figure 5. Application Launch Screen

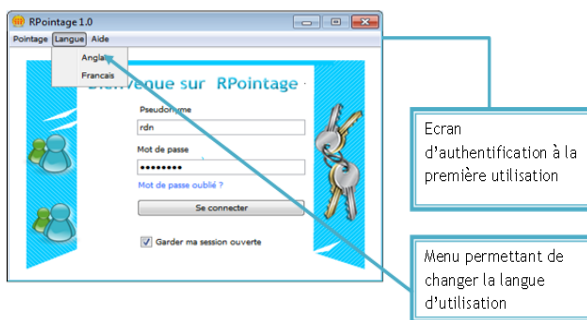


Figure 6. Authentication Screen

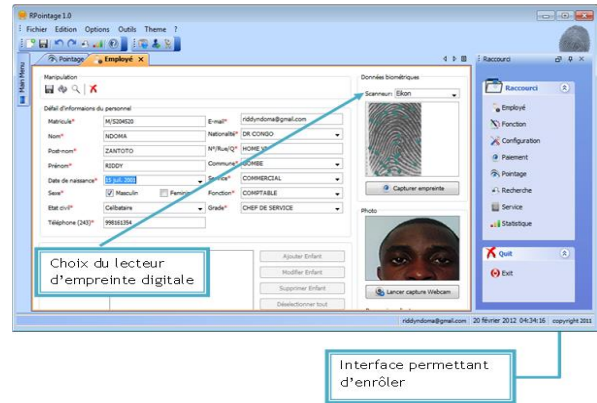


Figure 7. Enrollment Form

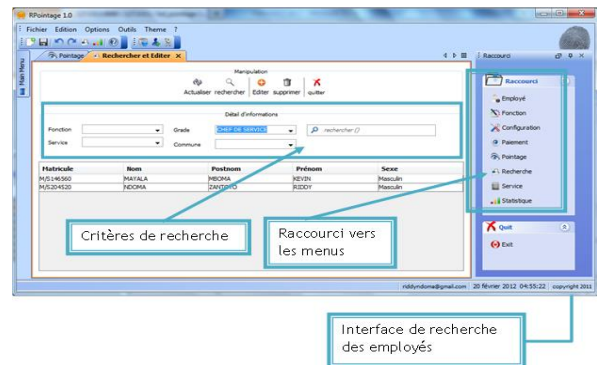


Figure 8. Search Form

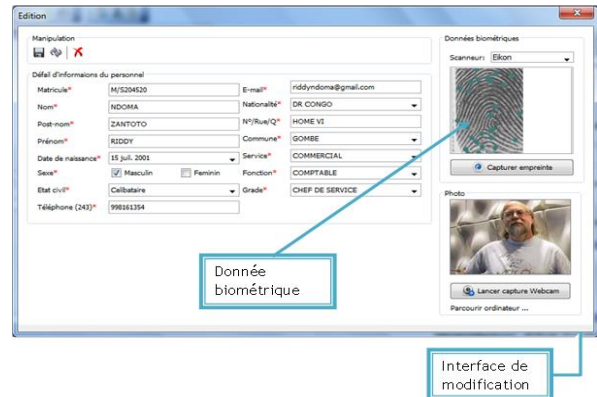


Figure 9. Change Form

II.6 SOURCES CODES EXTRACT

A: Database Scripts

```
-- Base de données: `bd_pointage`
-----
-- Structure de la table `t_commune`
CREATE TABLE IF NOT EXISTS `t_commune` (
  `id_commune` int(11) NOT NULL AUTO_INCREMENT,
  `nom_commune` text NOT NULL,
  PRIMARY KEY (`id_commune`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1
AUTO_INCREMENT=131341 ;
-----

-- Structure de la table `t_employe`
CREATE TABLE IF NOT EXISTS `t_employe` (
  `matricule` varchar(11) NOT NULL,
  `nom_employe` text NOT NULL,
  `postnom_employe` text NOT NULL,
  `prenom_employe` text NOT NULL,
  `date_naissance` date NOT NULL,
  `sexe_employe` text NOT NULL,
  `etat_civil` text NOT NULL,
  `nationalite` text NOT NULL,
  `phone_employe` int(11) NOT NULL,
  `mail_employe` text NOT NULL,
  `rue_employe` text NOT NULL,
  `photo_employe` longblob NOT NULL,
  `empreinte_employe` longblob NOT NULL,
  `id_commune` int(11) NOT NULL,
  `id_service` varchar(11) NOT NULL,
  `id_fonction` int(11) NOT NULL,
  `id_grade` int(11) NOT NULL,
  PRIMARY KEY (`matricule`),
  KEY `id_commune`
  (`id_commune`,`id_service`,`id_fonction`,`id_
  grade`),
  KEY `id_commune_2`
  (`id_commune`,`id_service`,`id_fonction`,`id_
  grade`),
  KEY `id_service` (`id_service`),
  KEY `id_fonction` (`id_fonction`),
  KEY `id_grade` (`id_grade`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;
-----

-- Structure de la table `t_enfant`
CREATE TABLE IF NOT EXISTS `t_enfant` (
  `id_enfant` int(11) NOT NULL AUTO_INCREMENT,
  `prenom_enfant` text NOT NULL,
  `nom_enfant` text NOT NULL,
  `date_naissance_enfant` date NOT NULL,
  `matricule` varchar(11) NOT NULL,
  PRIMARY KEY (`id_enfant`),
  KEY `matricule` (`matricule`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1
AUTO_INCREMENT=5 ;
-----

-- Structure de la table `t_fonction`
CREATE TABLE IF NOT EXISTS `t_fonction` (
  `id_fonction` int(11) NOT NULL,
  `nom_fonction` text NOT NULL,
  `salaire_base` double NOT NULL,
  `transport` double NOT NULL,
  `logement` double NOT NULL,
  `autres` double NOT NULL,
  PRIMARY KEY (`id_fonction`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;
-----

-- Structure de la table `t_grade`
CREATE TABLE IF NOT EXISTS `t_grade` (
  `id_grade` int(11) NOT NULL AUTO_INCREMENT,
  `nom_grade` text NOT NULL,
  PRIMARY KEY (`id_grade`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1
AUTO_INCREMENT=131370 ;
-----

-- Structure de la table `t_jourpointage`
CREATE TABLE IF NOT EXISTS `t_jourpointage` (
  `id_jourpointage` int(11) NOT NULL
  AUTO_INCREMENT,
  `date_jourpointage` date NOT NULL,
  PRIMARY KEY (`id_jourpointage`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1
AUTO_INCREMENT=11 ;
-----

-- Structure de la table
`
t_participerpointage`
CREATE TABLE IF NOT EXISTS
`t_participerpointage` (
  `matricule` varchar(11) NOT NULL,
  `id_pointage` int(11) NOT NULL,
  `heure_arrive` date NOT NULL,
  `heure_depart` date NOT NULL,
  PRIMARY KEY (`matricule`,`id_pointage`),
  KEY `id_pointage` (`id_pointage`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;
-----

-- Structure de la table `t_service`
CREATE TABLE IF NOT EXISTS `t_service` (
  `id_service` varchar(11) NOT NULL,
  `nom_service` text NOT NULL,
  PRIMARY KEY (`id_service`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;
-----

-- Structure de la table `t_user`
CREATE TABLE IF NOT EXISTS `t_user` (
  `id_user` int(11) NOT NULL AUTO_INCREMENT,
  `prenom` text NOT NULL,
  `nom` text NOT NULL,
  `postnom` text NOT NULL,
  `email` text NOT NULL,
  `nom_utilisateur` text NOT NULL,
  `mot_passe` text NOT NULL,
  `question` text NOT NULL,
  `reponse` text NOT NULL,
  `type` varchar(20) NOT NULL,
  `etat` varchar(20) NOT NULL,
  PRIMARY KEY (`id_user`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1
COMMENT='table contenant les utilisateurs'
AUTO_INCREMENT=4 ;
-- Contraintes pour les tables exportées
-- Contraintes pour la table `t_employe`
ALTER TABLE `t_employe`
ADD CONSTRAINT `t_employe_ibfk_1` FOREIGN KEY
(`id_commune`) REFERENCES `t_commune`
(`id_commune`) ON DELETE CASCADE ON UPDATE
CASCADE,
ADD CONSTRAINT `t_employe_ibfk_2` FOREIGN KEY
(`id_service`) REFERENCES `t_service`
(`id_service`) ON DELETE CASCADE ON UPDATE
CASCADE,
ADD CONSTRAINT `t_employe_ibfk_3` FOREIGN KEY
(`id_fonction`) REFERENCES `t_fonction`
(`id_fonction`) ON DELETE CASCADE ON UPDATE
CASCADE,
ADD CONSTRAINT `t_employe_ibfk_4` FOREIGN KEY
(`id_grade`) REFERENCES `t_grade`
(`id_grade`) ON DELETE CASCADE ON UPDATE
CASCADE;
-- Contraintes pour la table `t_enfant`
ALTER TABLE `t_enfant`
ADD CONSTRAINT `t_enfant_ibfk_1` FOREIGN KEY
(`matricule`) REFERENCES `t_employe`
```



```
(`matricule`) ON DELETE CASCADE ON UPDATE  
CASCADE;  
-- Contraintes pour la table  
`t_participerpointage`  
--TER TABLE `t_participerpointage`  
ADD CONSTRAINT `t_participerpointage_ibfk_2`  
FOREIGN KEY (`id_pointage`) REFERENCES  
`t_jourpointage` (`id_jourpointage`) ON  
DELETE CASCADE ON UPDATE CASCADE,  
ADD CONSTRAINT `t_participerpointage_ibfk_1`  
FOREIGN KEY (`matricule`) REFERENCES  
`t_employe` (`matricule`) ON DELETE CASCADE  
ON UPDATE CASCADE;
```

B : Application Scripts

To achieve our application, we had to write more than 2000 scripts. We do put here just a part for illustrative purposes.

```
packagepointage.docking;  
importcom.jidesoft.action.DefaultDockableBarD  
ockableHolder;  
importcom.jidesoft.docking.DockContext;  
importcom.jidesoft.docking.DockableFrame;  
importcom.jidesoft.docking.DockingManager;  
importcom.jidesoft.document.DocumentPane.Tabb  
edPaneCustomizer;  
importcom.jidesoft.utils.PortingUtils;  
  
public class MainFrame extends  
DefaultDockableBarDockableHolder {  
publicMainFrame(String title) throws  
HeadlessException {  
super(title);  
newConnexion();  
}  
publicMainFrame() throws HeadlessException {  
this("");  
}  
public static  
DefaultDockableBarDockableHoldershowFrame(fin  
al boolean exit) throws PropertyVetoException  
{  
_frame = new MainFrame("RPointage 1.0");  
_frame.setDefaultCloseOperation(JFrame.DISPOS  
E_ON_CLOSE);  
_frame.setIconImage(new  
ImageIcon(_frame.getClass().getResource("/poi  
ntage/ressources/note.png")).getImage());  
cbf = new MainFrameCommandBarFactory(_frame);  
// add a window listener to do clear up when  
windows closing.  
_windowListener = new WindowAdapter() {  
@Override  
public void windowClosing(WindowEvent e) {  
super.windowClosing(e);  
actionExit();  
}  
};  
_frame.addWindowListener(_windowListener);  
// set the profile key  
_frame.getLayoutPersistence().setProfileKey(P  
ROFILE_NAME);  
_frame.getLayoutPersistence().setXmlFormat(tr  
ue);  
// create tabbed-document interface and add  
it to workspace area  
_documentPane = createDocumentTabs();  
_documentPane.setTabbedPaneCustomizer(new  
TabbedPaneCustomizer() {  
@Override  
public void customize(final  
JideTabbedPanetabbedPane) {  
tabbedPane.setShowCloseButtonOnTab(true);  
tabbedPane.setShowCloseButtonOnSelectedTab(tr  
ue);
```

```
};  
});  
_frame.getDockingManager().getWorkspace().set  
Layout(new BorderLayout());  
_frame.getDockingManager().getWorkspace().add  
(_documentPane, BorderLayout.CENTER);  
_frame.getDockableBarManager().setProfileKey(  
PROFILE_NAME);  
// add toolbar  
_frame.getDockableBarManager().addDockableBar  
(MainFrameCommandBarFactory.createMenuComm  
andBar());  
_frame.getDockableBarManager().addDockableBar  
(MainFrameCommandBarFactory.createStandardCom  
mandBar());  
//  
_frame.getDockableBarManager().addDockableBar  
(MainFrameCommandBarFactory.createOptionsComm  
andBar());  
_frame.getDockableBarManager().addDockableBar  
(MainFrameCommandBarFactory.createToolsComm  
andBar());  
// add status bar  
_statusBar = createStatusBar();  
_frame.getContentPane().add(_statusBar,  
BorderLayout.AFTER_LAST_LINE);  
_frame.getDockingManager().getWorkspace().set  
AdjustOpacityOnFly(true);  
_frame.getDockingManager().setUndoLimit(10);  
_frame.getDockingManager().beginLoadLayoutDat  
a();  
// add all dockable frames  
_frame.getDockingManager().addFrame(MainFrame  
CommandBarFactory.createFramePrincipal(_frame  
));  
_frame.getDockingManager().addFrame(MainFrame  
CommandBarFactory.createFrameRaccourci());  
_frame.getDockingManager().setShowGripper(tru  
e);  
_frame.getDockingManager().setOutlineMode(Doc  
kingManager.TRANSPARENT_OUTLINE_MODE);  
_frame.getDockingManager().setPopupMenuCustom  
izer(new  
com.jidesoft.docking.PopupMenuCustomizer() {  
@Override  
public void customizePopupMenu(JPopupMenu  
menu, final DockingManagerdockingManager,  
final DockableFramedockableFrame,  
booleanonTab) {  
menu.addSeparator();  
menu.add(new AbstractAction("Move to Document  
Area") {  
@Override  
public void actionPerformed(ActionEvent e) {  
dockingManager.removeFrame(dockableFrame.getK  
ey(), true);  
DocumentComponentdocumentComponent = new  
DocumentComponent((JComponent)  
dockableFrame.getContentPane(),  
dockableFrame.getKey(),  
dockableFrame.getTitle(),  
dockableFrame.getFrameIcon());  
_documentPane.openDocument(documentComponent)  
;  
_documentPane.setActiveDocument(documentCompo  
nent.getName());  
}  
});  
}  
});  
// load layout information from previous  
session  
_frame.getLayoutPersistence().loadLayoutData(  
);  
if (Lm.DEMO) {  
Lm.z();  
}
```

```
_frame.toFront();
return _frame;
}
public static void actionResearch() {
if
(!_documentPane.isDocumentOpened("Rechercher
et Editer")) {
try {
_documentPane.openDocument(new
DocumentComponent(new SearchEdit(_frame),
"Rechercher et Editer", "Rechercher et
Editer"
+ "", new
ImageIcon(_frame.getClass().getResource("/ima
ges/gnome-orcal2.png")));
} catch (Exception ex) {
ex.printStackTrace();
}
}
_documentPane.setActiveDocument("Rechercher
et Editer");
}
private static void clearUp() {
_frame.removeWindowListener(_windowListener);
_windowListener = null;
if (_frame.getLayoutPersistence() != null) {
_frame.getLayoutPersistence().saveLayoutData(
);
}
if (_documentPane != null) {
_documentPane.dispose();
_documentPane = null;
}
if (_statusBar != null &&
_statusBar.getParent() != null) {
_statusBar.getParent().remove(_statusBar);
}
_statusBar = null;
_frame.dispose();
_frame = null;
}
public static StatusBarcreateStatusBar() {
// setup status bar
StatusBarstatusBar = new StatusBar();
progress = new ProgressStatusBarItem();
statusBar.add(progress, JideBoxLayout.VARY);
finalLabelStatusBarItem label = new
LabelStatusBarItem("");
label.setText("riddyndoma@gmail.com");
label.setAlignment(JLabel.CENTER);
statusBar.add(label, JideBoxLayout.FLEXIBLE);
finalTimeStatusBarItem time = new
TimeStatusBarItem();
time.setTextFormat(new SimpleDateFormat("dd
MMM yyyyHH:mm:ss"));
statusBar.add(time, JideBoxLayout.FLEXIBLE);
// elements added on Statusbar by RDN
finalLabelStatusBarItemmyJL = new
LabelStatusBarItem("");
myJL.setText("copyright 2011");
myJL.setAlignment(JLabel.CENTER);
statusBar.add(myJL, JideBoxLayout.FLEXIBLE);
// Font added on Statusbar by RDN
Font myFont = new Font("tahoma", Font.PLAIN,
10);
statusBar.setFont(myFont);
returnstatusBar;
}
private static
DocumentPanecreateDocumentTabs() {
DocumentPanedocumentPane = new DocumentPane()
{
// add function to maximize (autohideAll) the
document pane when mouse double clicks on the
tabs of DocumentPane.
@Override
protectedIDocumentGroupcreateDocumentGroup()
{
IDocumentGroup group =
super.createDocumentGroup();
if (group instanceofJideTabbedPane) {
((JideTabbedPane) group).addMouseListener(new
MouseAdapter() {
@Override
public void mouseClicked(MouseEvent e) {
if (SwingUtilities.isLeftMouseButton(e)
&&e.getClickCount() == 2) {
if (!_autohideAll) {
_fullScreenLayout =
_frame.getDockingManager().getLayoutRawData(
);
_frame.getDockingManager().autohideAll();
_autohideAll = true;
} else {
if (_fullScreenLayout != null) {
_frame.getDockingManager().setLayoutRawData(
_fullScreenLayout);
}
_autohideAll = false;
}
}
Component lastFocusedComponent =
_documentPane.getActiveDocument().getLastFocu
sedComponent();
if (lastFocusedComponent != null) {
lastFocusedComponent.requestFocusInWindow();
}
}
});
return group;
}
};
documentPane.setTabPlacement(JTabbedPane.TOP)
;
documentPane.setPopupMenuCustomizer(new
PopupMenuCustomizer() {
@Override
public void customizePopupMenu(JPopupMenu
menu, final IDocumentPane pane, final String
dragComponentName,
finalIDocumentGroupdropGroup, booleanonTab) {
if
(!pane.isDocumentFloating(dragComponentName))
{
menu.addSeparator();
menu.add(new AbstractAction("Dock to the
Side") {
@Override
public void actionPerformed(ActionEvent e) {
DocumentComponentdocumentComponent =
pane.getDocument(dragComponentName);
if (documentComponent != null) {
pane.closeDocument(dragComponentName);
// check if the document is really closed.
There are cases a document is not closable or
veto closing happens which can keep the
document open after closeDocument call.
if
(!pane.isDocumentOpened(dragComponentName)) {
finalDockableFrame frame = new
DockableFrame(documentComponent.getName(),
documentComponent.getIcon());
frame.setTabTitle(documentComponent.getTitle(
));
frame.getContentPane().add(documentComponent.
getComponent());
frame.setInitIndex(0);
frame.setInitSide(DockContext.DOCK_SIDE_EAST)
;
frame.setInitMode(DockContext.STATE_FRAMEDOCK
ED);
_frame.getDockingManager().addFrame(frame);

```

```
_frame.getDockingManager().activateFrame(frame.getKey());
}
}
});
return documentPane;
}
private static
JComponent createMultiViewDocument(String
fileName) {
JideTabbedPane pane = new
JideTabbedPane(JideTabbedPane.BOTTOM);
pane.setTabShape(JideTabbedPane.SHAPE_BOX);
pane.addTab("Design",
createTextArea(fileName));
pane.addTab("HTML",
createTextArea(fileName));
return pane;
}
private static
JComponent createTextArea(String fileName) {
JTextArea area = new JTextArea();
Document doc = new PlainDocument();
try {
InputStream in =
MainFrame.class.getResourceAsStream(fileName);
if (in == null) {
in = new FileInputStream(fileName);
}
byte[] buff = new byte[4096];
int nch;
while ((nch = in.read(buff, 0, buff.length))
!= -1) {
doc.insertString(doc.getLength(), new
String(buff, 0, nch), null);
}
area.setDocument(doc);
} catch (IOException | BadLocationException
e) {
System.out.println(e.getLocalizedMessage());
}
return area;
}
@Override
protected ContentContainer createContentContain
er() {
return new LogoContentContainer();
}
private class LogoContentContainer extends
ContentContainer {
@Override
protected void paintComponent(Graphics g) {
super.paintComponent(g);
ImageIcon imageIcon = new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/empreinte.png"));
imageIcon.paintIcon(this, g, getWidth() -
imageIcon.getIconWidth() - 2, 2);
}
}
public static void actionNewChild() {
if
(!_documentPane.isDocumentOpened("Enfant")) {
_documentPane.openDocument(new
DocumentComponent(new NewChild(_frame),
"Enfant", "Enfant"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_child.png"))));
}
_documentPane.setActiveDocument("Enfant");
}
public static void actionNewPartener() {
if
(!_documentPane.isDocumentOpened("Conjoint"))
{
_documentPane.openDocument(new
DocumentComponent(new NewPartener(_frame),
"Conjoint", "Conjoint"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_partener.png"))));
}
_documentPane.setActiveDocument("Conjoint");
}
}
public static void actionNewWorker() {
if
(!_documentPane.isDocumentOpened("Employé"))
{
try {
_documentPane.openDocument(new
DocumentComponent(new NewWorker(_frame),
"Employé", "Employé"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_worker.png"))));
} catch (SQLException ex) {
ex.printStackTrace();
}
}
_documentPane.setActiveDocument("Employé");
}
}
public static void actionNewService() {
if
(!_documentPane.isDocumentOpened("Service"))
{
_documentPane.openDocument(new
DocumentComponent(new NewService(_frame),
"Service", "Service"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_service.png"))));
}
_documentPane.setActiveDocument("Service");
}
}
public static void actionNewStag() {
if
(!_documentPane.isDocumentOpened("Stagiaire"))
{
_documentPane.openDocument(new
DocumentComponent(new NewStag(_frame),
"Stagiaire", "Stagiaire"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_stag.png"))));
}
_documentPane.setActiveDocument("Stagiaire");
}
}
public static void actionNewCommune() {
if
(!_documentPane.isDocumentOpened("Commune"))
{
_documentPane.openDocument(new
DocumentComponent(new NewCommune(_frame),
"Commune", "Commune"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_service.png"))));
}
_documentPane.setActiveDocument("Commune");
}
}
public static void actionNewGrade() {
if (!_documentPane.isDocumentOpened("Grade"))
{
_documentPane.openDocument(new
DocumentComponent(new NewGrade(_frame),
"Grade", "Grade"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_service.png"))));
}
_documentPane.setActiveDocument("Grade");
}
}
public static void actionNewFonction() {
```

```
if
(!_documentPane.isDocumentOpened("Fonction"))
{
_documentPane.openDocument(new
DocumentComponent(new NewFonction(_frame),
"Fonction", "Fonction"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/color-line.png"))));
}
_documentPane.setActiveDocument("Fonction");
}
public static void actionNewMatching() {
if
(!_documentPane.isDocumentOpened("Pointage"))
{
_documentPane.openDocument(new
DocumentComponent(new NewMatching(_frame),
"Pointage", "Pointage"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/14_layer_novisible.png"))));
}
_documentPane.setActiveDocument("Pointage");
}
public static void actionSetting() {
}
public static void actionStatistic() {
if
(!_documentPane.isDocumentOpened("Statistique
")) {
_documentPane.openDocument(new
DocumentComponent(new Stat(_frame),
"Statistique", "Statistique"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_service.png"))));
}
_documentPane.setActiveDocument("Statistique"
);
}
public static void actionPaie() {
if
(!_documentPane.isDocumentOpened("Paie"))
{
_documentPane.openDocument(new
DocumentComponent(new GestPaie(_frame),
"Paie", "Paie"+ "", new
ImageIcon(_frame.getClass().getResource("/poi
ntage/ressources/mini_service.png"))));
}
_documentPane.setActiveDocument("Paie");
}
public static void actionNewPerson() throws
UnsupportedLookAndFeelException {
```

General Conclusion

We made an automatic pointing system by fingerprint based on the extraction of minutiae. We obtained interesting results which allow a better accuracy for agents' payroll by eliminating calculation errors, bad adjustments when rounded, involuntary overpaid waste of time. It also provides statistics on all of the personnel on the basis of well-defined criteria. The indexing of a database in this application confirms the need for tomorrow's biometrics.

References

Books, articles, lecture's notes

[ABD1] D. Abdelhakim, «Approches bio-inspirées pour la reconnaissance des formes», mémoire de fin d'études, Université MENTOURI DE CONSTANTINE, 2006

[BAT2] J.D. BATUBENGA « Etude d'un système de reconnaissance des formes biométriques Application sur l'identification des empreintes digitales», mémoire de DES d'études, Université de Kinshasa, 2010.

[BOU3] F. BOUGAMOUZA, «Contribution à la reconnaissance automatique de l'écriture manuscrite arabe, application sur les montants littéraux des chèques», thèse de doctorat, 2009

[CHA4] A. CHAARI, « Nouvelle approche d'identification dans les bases de données biométriques basée sur une classification non supervisée », thèse, octobre 2009.

[CHA5] F. CHAFIA, « Un Crypto système à Base de la Biométrie Pour l'Authentification », mémoire de fin d'études, 2010.

[CHA6] B. CHARROUX, A. OSMANI & Y. THIERRY-MIEG, «UML 2 pratique de la modélisation», Edition Pearson 2eme Edition, 2009.

[DJU7] S. DJUNGU, «Cours Génie Logiciel et Construction des Programmes», Unikin, Notes de cours 2010.

[GAB8] J. GABAY & D. GABAY, «UML 2 Analyse et conception», Edition Dunod, 2008.

[MOR9] N. MORIZET, « Reconnaissance Biométrique par Fusion Multimodale du Visage et de l'Iris », thèse, École nationale d'Informatique et Télécommunications de Paris, mars 2009.

[MBU10] E. MBUYI MUKENDI, «Cours de Base de données», Unikin, Notes de cours 2009.

[PER11] F. PERRONNIN et J. DUGELAY, « Introduction à la Biométrie Authentification des Individus par Traitement Audio-Vidéo », Revue Traitement du Signal, volume 19, 2002.

[ROQ12] P. Roques & F. Vallée, «UML en Action», Edition Eyrolles 4eme Edition, 2007.

SAPWOOD OAK CLASSIFICATION USING GLOBAL AND LOCAL FEATURES

Hammad Naeem¹, Asim Anees², Maria Minhas³ and Jameel Ahmed⁴

^{1,4}Department of Electrical Engineering ,Hitec University, Taxila Cantt,47040,Pakistan.

²Department of Environmental Engineering, University of Tasmania, Australia,

³Department of Electrical Engineering,University of Minnesota, Duluth, 55812,United states.

Abstract

Classification of wood images has always fascinated researchers and is a hot area of research as it brings automation in wood industry. Wood comes in different types and for a given wood; it can have some good part and some bad part like cracks, water affected and sapwood. The objective of this paper is to classify the sapwood region and the non-sapwood region in oak wood image in order to get the good part of the wood (heart part). Color based features and texture based feature extraction strategies were employed and it was found that color based feature extraction techniques are better compared to others in order to find sapwood.

Keywords: SAPWOOD detection, Feature Combination, Gabor Filter.

1. Introduction

Oak-sapwood classification is one of interesting subject in wood classification. Sapwood is part of the wood which is categorized as bad part because it is less strong than the heart part. Distinguishing between sapwood and heart part of the wood is the main focus of this paper. Several difficulties were faced in setting criteria for classification of sapwood region, for example, sapwood color is hardly uniquely described because some of them are different among different woods. The coloration of the heartwood and sapwood can be sharply contrasted with each other, but the coloration can be same for both heartwood and sapwood in some other trees. This can be caused by the age of tree or environment of growth.

In order to recognize sapwood, one needs to know the characteristic of sapwood. Sapwood is the cells within the wood that is used to transfer water and minerals from the roots to the leaves store the reserved food materials and provide strength to the whole tree [1].

The system starts with an image acquisition process where the input samples are captured using a camera (This process done by Luxscan).

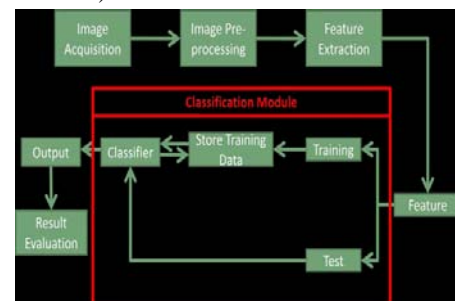


Fig 1: Flow Diagram of Process

After the input image is captured, it will be normalized through the pre-processing process. The normalized images will be used for feature extraction and important textural features are extracted here. The extracted features will be trained using our ground truth and then classified using a classifier and produces an output as shown in Figure 1.



Fig 2: Ground Truth Images

Pre-processing is applied on the input images in order to eliminate no-information area in the wood. The black background from the image seems has no information in it. It should be removed to reduce the input size and to enhance identifying the characteristic of the wood so that it will help in the recognition process. The ground truth images are formed as shown in figure 2.

2. Feature Extraction

Several different techniques were employed and these techniques are explained as follows

2.1 Color Feature

Color features consist of hue, illumination and chromacity (saturation). Hue refers to color attribute. The luminosity is the light component of a color which is from white to black. The saturation is the intensity of the color from grey to color. There are various color space available such as RGB (Red, Green, Blue), HSV (Hue, Saturation, brightness Value) and CMY (Cyan, Magenta, Yellow) etc.

2.2 Relative Mean Intensity

In maximum type of woods, sapwood have different color from heartwood. Heartwood is considered to have the same intensities as the mean intensities of the full image [2]. Meanwhile, almost all of the sapwood has brighter intensities than heartwood. Only few of them have darker intensities. At first, it is also considered as the dark-sapwood so that absolute relative mean can be used as the feature extraction. As shown in Figure 3, sapwood are darker in that wood so that when we use the relative mean intensity features, it gives brighter heartwood and dark sapwood wood (Figure 3b). If we use the absolute relative mean features (Figure 3c), sapwood and heartwood are bright, we cannot distinguish them well.

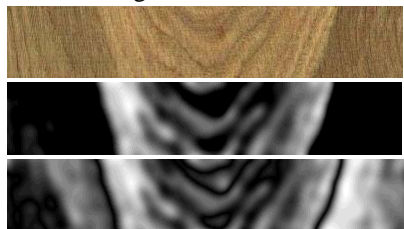


Fig 3: a) Original Image b) Relative Mean Intensity Image c) Absolute Mean Image

Results show that when Features are extracted using absolute relative mean, could not classify the sapwood well. Moreover, other defects and cracks are also classified as sapwoods because of their lower intensities. Therefore, relative mean intensity is chosen as the feature extraction because it can give a better separation between sapwood and heartwood. The formula is defined as:

$$\text{Features} = \text{Pixel intensity} - \text{Mean global image intensity} \dots \dots \dots (1)$$

Pre-processing is done to get a more compact classified sapwood region. Gaussian smoothing is used because it is good to reduce noise and to blur the images[3]. Parameters that are tuned are the kernel size and standard deviation. The size of the kernel should normally be selected large enough and increasing the standard deviation will increase the blurring.

2.3 Gabor Filter

Gabor filter is also known as Gabor wavelets. A Gabor function is a complex sinusoid modulated by a rotated Gaussian. A two dimensional Gabor function consists of a sinusoidal plane wave of some frequency and orientation, modulated by a two-dimensional Gaussian.

$$g_e(x, y) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right) e^{-\left(\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right)} \cos((2\pi(x/\lambda) + \psi)) \dots (2)$$

Where,

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \dots \dots \dots (3) \\ y' &= x \sin \theta - y \cos \theta \dots \dots \dots (4) \end{aligned}$$

In the above equation, λ represents the wavelength of the cosine factor, θ represents the orientation of the normal of a Gabor function in degrees, ψ is the phase offset in degrees, and γ is the spatial aspect ratio and it specifies the elliptically of the Gabor function, and σ is the standard deviation of the Gaussian.

Gabor filters have various properties that make them particularly suitable for texture segmentation [4]. It has been shown that the Gabor function is a band-pass filter that can be tuned to a narrow set of frequency anywhere in the frequency domain. The output of properly parameterized Gabor channels can thus be used to reconstruct the most important features of a textured image.

2.4 GLDM

Conventional GLCM method is two dimensional as it focus on the co-occurrence of the specific pixel pair, so it is much time consuming. Therefore, GLDM as 1-dimensional GLCM can be a solution to answer this problem.

To reduce computations, the GLCM dimension can be reduced from two dimensions to one dimension by combining certain values of the matrix [5]. By focusing only on the differences of the grey level, we are only concerning on a one-dimensional GLCM with a significantly smaller size which is only $2 \times G - 1$ where G represents the grey level, compared to $G \times G$ for a conventional two-dimensional GLCM.

GLDM uses the joint probability density function which normalizes the GLCM by dividing every set of pixel pairs with the total number of pixel pairs used. GLDM is based on the occurrence of two pixels which have a given absolute difference in gray level and separated by a specific displacement vectors. The same texture measurements used in GLCM can also be applied for GLDM, but it must be modified to suite the one dimensional computation. In implementation, the GLDM is seen as the density function so it is easier to calculate the statistical values from it.

3. Implementation and Results

In this section, implementation of different techniques is discussed.

3.1 Relative Mean Intensity

Relative mean intensity features are extracted from images. The results in figure 4 show sapwood have much higher values than the mean and can be characterized well using this method. In order to avoid the sparse and noisy sapwood classification results, we applied the Gaussian smoothing. For the parameters, we set the kernel size to be 25 and standard deviation is 5 to achieve enough blurring and uniform sapwood region.

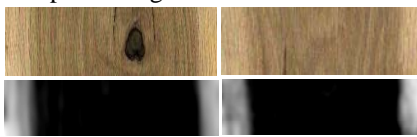


Fig 4: Result of Relative Mean Intensity feature extraction from wood patches

3.2 Gabor Filter

To analyze how Gabor filter segment the wood image texture, result was observed by convolving the Gabor filter via FFT in different orientations and scales. First, it was tested with the patches having sapwood only and the non-sapwood patches only to see how Gabor filter can give a measurements are seen in Figure 5. By observation, entropy always gives a good separation between sapwood and heartwood. In those two patches, sapwood is defined to have high entropy values. Contrast is also a good feature but it is not too obvious in characterize the sapwood. Energy is the inverse of the entropy but it is not too clear. Homogeneity fails to show the difference between sapwood and heartwood region. Therefore, for simplicity, we only use the entropy feature extracted from GLDM. It gives the high value for sapwood region (red), even though the color is not brighter than the heartwood.

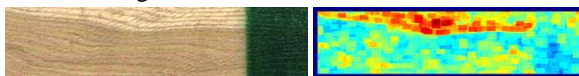


Fig 6: Entropy feature extraction from GLDM

3.3 Relative Mean Intensity + Gabor features and ANN

Gabor features is combined with relative mean features. For Gabor feature, we select the scale and orientation, which is at scale = 3 and first orientation (0°). Some of

specific characteristic for sapwood. We used 5 sapwood patches and 5 non-sapwood patches from 5 training images. Unfortunately, we could not infer any separation between texture of sapwood and non-sapwood in different orientations and scales. On the other hand, although in some patches the non-sapwood have more random texture, but some Gabor features of heartwood may also have the same texture characteristic with sapwood. This makes the criteria of sapwood texture not clearly defined by Gabor.

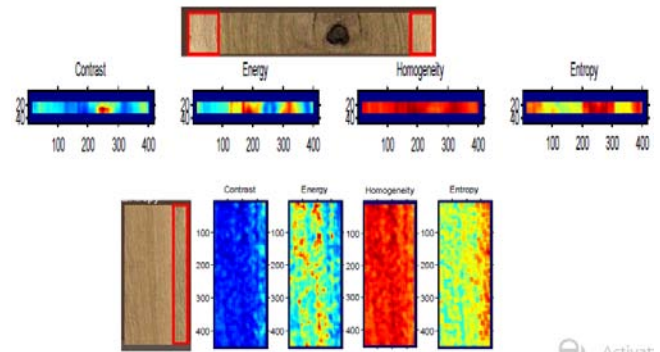


Fig 5: Image textures extracted from GLDM statistical measurements Direction of orientation is selected to be 0 deg, since the most of the sapwoods in our sets has uniform vertical pattern. If we choose smaller distance than 5, it could not give a clear separation while if we choose a bigger distances the result even worse. After, GLDM is applied in several image patches that contain sapwood, image textures of the four statistical the classifier output using this combined feature is shown in the figure 7.



Fig 7: CLASSIFICATION RESULTS using Gabor + RELATIVE MEAN FEATURES

Visually, it seems worse since it gives more false positive classification of sapwoods. The classifier output with Gabor and relative mean features give lower classification performance. It is shown in the parameters derived from confusion matrix, as the following: Accuracy = 0.8992, TPR = 0.2470, FPR = 0.0311, F_score = 0.0896.

3.4 Relative Mean Intensity + GLDM and ANN

GLDM is combined with relative mean features so that the classifier can have input from color as well as the texture features. The performance measures from confusion matrix are as follows: Accuracy = 0.95, TPR = 0.61, FPR = 0.0169, F_score = 0.372. From these measures, it can be seen that actually the network only gives TPR 0.61. Otherwise, it is tried to use these combined GLDM and Relative Mean

features to the testing images. Results are shown in figure 8.



Fig 8: Classification results for GLDM + relative mean features

4. Conclusion

Sapwood and heartwood can be classified well in our strategies, by using the appropriate combination of features extraction method and classifiers. The best features extraction method for supervised and unsupervised classification is the relative mean intensity and GLDM. Ground truth images are created but with imperfect masking which may lead to a reduced performance in classification. Whereas in unsupervised, the evaluation is based on visual assessment of sapwood region. To summarize, we conclude that classifiers can work well if we have good training set and the sapwood has a uniform criteria in most of the cases.

References

- [1] Z. Zhang, et al. Locating the Wood Defects with Typical Features and SVM. 2008. School of Information Technology, NanJing Forestry University.
- [2] J.Yi Tou. Texture classification on wood images for species recognition. December 2009. Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman Malaysia.
- [3] 'The colour Image processing Handbook' by S.J Sangwine and R.E.N Horne, Chapman & Hall.
- [4] 'Fundamentals of Computer Vision' by Mubarak Shah, University of Central Florida, Orlando
- [5] 'Texture analysis aids interpretation' by Satinder Chopra and Vladimir Alexeev, Arcis Corp.,Calgar

Hammad Naeem received his degree of Master in Computer Vision and Robotics from Heriot Watt University-Edinburgh, University of Girona-Spain and University of Burgundy-France. He was on an Erasmus Mundus program which was funded by European Union. He holds Bachelors in Electrical Engineering from Air University, Islamabad,Pakistan. Currently he is serving as a Lecturer in Department of Electrical Engineering at HITEC University Taxila, Pakistan. He has particular interests in Vision based robotics and image processing based intelligent machines.

Asim Anees has done Master's in Computer Vision and robotics. He was funded by European Union for his Master Degree under the scholarship scheme named Erasmus Mundus (Vibot). He holds a Bachelor degree in Electrical Engineering from UET, Peshawar, Pakistan. Currently he is pursuing his Phd degree from University of Tasmania, Australia.

Maria Minhas is currently doing her Master's in Electrical Engineering from University of Minnesota, Duluth, USA. She did her Bachelor in Electrical Engineering from HITEC University, Taxila, Pakistan in 2011. She has research interest in image processing and Bio-medical vision based solutions.

Jameel Ahmed received BSc. degree in Electronic Engineering from NED University of Engineering & Technology, Karachi, Pakistan in 1991. He did MSc. Electrical Engineering (Telecommunication) from NUST, Pakistan and PhD (Telecommunication Engineering) from Hamdard University Pakistan and Nanyang Technological University (NTU) Singapore. Subsequently, carried out Post Doctoral Fellowship from NTU Singapore. Jameel Ahmed is actively involved in teaching and research for the last twenty years. Presently, he is serving as professor and chairman, department of Electrical Engineering at HITEC University, Taxila Cantt-Islamabad, Pakistan. Dr Jameel is also heading the Signal Processing & Communication Systems (SPCS) research group of the department.

A Comparative Study about Object Classification Based On Global and Local Features

Hammad Naeem¹, Maria Minhas² and Jameel Ahmed³

¹Department of Electrical Engineering ,HITEC University, Taxila Cantt,47040,Pakistan.

²Department of Electrical Engineering, University of Minnesota, Duluth, 55812,United states.

³Department of Electrical Engineering , HITEC University, Taxila Cantt,47040,Pakistan.

Abstract

Scene classification and object recognition is a hot area of research in the field of computer vision and has always fascinated researchers to explore strategies for optimization of results. Global and local features are manipulated to find a match in the images or scene categories. This paper mainly comprises of finding the scene labels based on the objects present in it. The image is transformed into a feature space and the classifier is trained to differentiate each class in the feature space. Various feature extraction techniques like RGB histogram , SIFT and covariance are explored in this paper to find an optimized result. Different classifiers were tested individually as well as their combinations to achieve better results. Combination of Sparse SIFT and Dense SIFT techniques was found to perform better compared to others.

Keywords: Scene Classification, object recognition, Bag of words, Sparse SIFT, Dense SIFT.

1. Introduction

Object recognition in an image is an old yet inconclusive area of research in computer vision. The major problems encountered in object recognition are objects present in the scene having high clutter, illumination variations, high occlusion, high degree of geometric transformations and intra class variation.

Early techniques were mainly appearance based that is using the global features of the image. Origin is the empirical appearance based technique which involves sub-space methods and histograms. Their major drawback was requirement of large number of training images yielding a high computational cost [1]. Also as global features are usually sensitive to variations; these methods were unable to tolerate the geometrical transformations.

So, these methods were replaced by techniques which used local invariant features. Since local features are more robust to clutter, occlusion, light changes and geometric transformations hence, strategies using them proved to be more successful [1]. Local features were successfully utilized for a long time but the problem with local features was that they normally lose all the information about the spatial layout of the image and hence lack the descriptive ability. However it should be mentioned that local features are playing an important role in classification and recognition tasks.

In this paper, we have worked on the PASCAL dataset from 2006 in which there are ten different classes of object and we have to detect a particular object in a given image. The ten different object classes are: bicycle, bus, car, cat, cow, dog, horse, motorcycle, people and sheep. Classification is not easy as objects have high degree of geometrical transformation, illumination changes, clutter, occlusion and intra-class variations [1]. There are also cases in which an image contains objects from different classes. This only contributes in making classifier confused in deciding about the particular object in the image.

2. General Strategy

Fairly, large set of training images is provided for each of the ten classes. Classifier is trained with the help of these training images. First features are extracted from images of each class using any feature extraction technique and are then mapped according to the bag of words vocabulary. The feature spaces generated by extraction of features from the images of different classes are grouped into k clusters with the help of k-mean algorithm. The numbers of clusters generated define the size of vocabulary. This technique of grouping or clustering of features is called 'Bag of words' technique.

Bag of words technique has demonstrated a fair deal of success and has surpassed many past techniques. Next histogram is generated for each of the bag of words feature and classifier is trained for each of the class. The classifier is trained to indicate the presence or absence of an object in the image with some level of confidence. Classifier is trained with positive and negative examples of the class to generate decision boundaries for a class in the feature space.

Ground truths are provided for each image. They are compared with the labels assigned by the classifier and ROC curves are generated showing the performance of technique for that particular class.

It is necessary to mention that histograms are normalized in order to have same number of features from each image. For maximum computational efficiency, we normalize all histograms by the total weight of all features in the image, in effect forcing the total number of features in all images to be the same.

2.1 RGB Histogram

First strategy used was computing the histogram of the image. Training images are separated into R, G and B components and histograms are generated. This technique is tested to check the performance of semantic information of an image in object detection. It is expected to give good results with object classes of almost similar background all the time for example car which is expected to have a road (black or grey) or greenery in the background majority of time. Similarly cow might have grass field (green) or sky (blue) background for most of times.

2.2 SIFT (Scale Invariant Feature Transform)

Next strategy is based on the local invariant features. SIFT (Scale Invariant Feature Transform) has brought revolution in the field of computer vision. As SIFT features are fairly robust to geometric variations and perform better in the presence of occlusion and clutter [3], they provide very good features for object or scene labeling. The SIFT feature can be extracted using two techniques, Sparse SIFT and Dense SIFT.

Sparse SIFT:In this, only strong SIFT features are extracted from the image. Instead of taking features from every part of the image uniformly, sparse sift only returns feature which are strong and distinct. There are two or three factors which affect the performance of the sparse SIFT. The most important one is the threshold value used in the SIFT function. The threshold value decides which peaks have to be filtered or rejected in extracting SIFT features. High value of this threshold means that the number of features extracted will be less and some of the features lie below the threshold specified. We tried different values of the threshold and the findings are discussed in the results section.

Dense SIFT:In this technique, SIFT features are uniformly extracted from every part of an image. In this method,

window frames with specific scales are defined and then whole image is scanned to extract features for every window frame. This results in a large number of features as now we not only extract strong features but also some weak features. So we used different values of the scale for the frames and received different number of features as discussed in the results section.

2.3 Covariance

It is always a good idea to find the covariance within the image. This is useful in the sense that a particular feature or object has some covariance with another object or a particular background in the image for example cow might have some covariance with the green or blue regions depicting grass or sky in the background. A feature image is constructed where each pixel is represented by the vector of “d” pixel level features [4].Hence any region can be represented by ad-by-d covariance matrix. For comparing two covariance matrices, Tuzet. al. devised a distance metric based on their generalized eigenvectors[4].The covariance matrix is full-rank, so there will be 9 eigenvectors for a feature vector of length 81.The relation is given in following equation.

$$F(x, y)=[x, y, R(x, y), G(x, y), B(x, y), \left| \frac{\partial I(x, y)}{\partial x} \right|, \left| \frac{\partial I(x, y)}{\partial y} \right|, \left| \frac{\partial^2 I(x, y)}{\partial x^2} \right|, \left| \frac{\partial^2 I(x, y)}{\partial y^2} \right|] \tag{1}$$

3. Results

Results for the various strategies used areas follows.

3.1 RGB Histograms

As a start, the trial was done with the RGB histogram feature of the images. The results are shown in table 1.

Table 1: RGB histogram feature of the images

Object	Accuracy
Bicycle	0.401
Bus	0.467
Car	0.613
Cat	0.592
Cow	0.706
Dog	0.524
Horse	0.628
Motorbike	0.563
Person	0.515
Sheep	0.647
Average	0.576

Low Accuracies were expected because object categories with a lot of variation in the background, will not be classified well using RGB features. Best results were obtained for the Cow as expected because for most of the cow scenes, there is usually a greenery and sky in the background, so RGB histograms are able to model these images in a good way. Hence, it can be concluded that RGB histogram can perform good if there is a high correlation between the object type and colour of the background

3.2 SIFT (Scale Invariant Feature Transformation)

SIFT features were tested with a number of parameters which are discussed in the following sections:

3.2.1 Vocabulary Parameters

As for SIFT, ‘Bag of Features’ technique was used, so three different parameters were tested for the generation of the vocabulary. The parameters include number of images used for generating the vocabulary, size of the vocabulary (number of clusters or words in the vocabulary) and threshold value for extracting the SIFT features from the images. These are explained in the following sections a, b and c.

(a) Number of Images: For this, the number of images used to generate the vocabulary is varied. The other two parameters, vocabulary size and threshold were kept constant at ‘500’ and ‘0’. The numbers of images used are 5, 10 and 15. The results are shown in table 2.

Table 2: Results for varying number of images

	5	10	15
Bicycle	0.611	0.625	0.632
Bus	0.392	0.365	0.365
Car	0.665	0.685	0.673
Cat	0.628	0.649	0.655
Cow	0.638	0.673	0.679
Dog	0.486	0.493	0.471
Horse	0.401	0.394	0.426
Motorbike	0.568	0.515	0.508
Person	0.621	0.628	0.628
Sheep	0.717	0.725	0.711

It can be observed from table 2 that by increasing the number of images per class for vocabulary generation improved the results. However, increasing the number of images does not guarantee an improvement in the results. Hence, no definitive conclusion can be drawn between the performance improvement and number of images used to generate vocabulary. It might be because of the quality of images used for that particular class.

b) Varying Vocabulary Size: In this, the vocabulary size was varied while keeping the threshold and number of images constant at ‘0’ and ‘10’ respectively. The value of ‘k’ was varied in K-mean algorithm, which is used to cluster features to generate different sizes for vocabulary. The values of ‘k’ taken were 200, 500 and 1000. Thus we get 200, 500 and 1000 clusters or words in the vocabulary. The results are shown in table 3.

Table 3: Results for varying the vocabulary size

	200	500	1000
Bicycle	0.649	0.625	0.775
Bus	0.498	0.365	0.667
Car	0.728	0.685	0.745
Cat	0.495	0.649	0.614
Cow	0.701	0.673	0.815
Dog	0.496	0.493	0.549
Horse	0.452	0.394	0.400
Motorbike	0.542	0.515	0.573
Person	0.494	0.628	0.539
Sheep	0.607	0.725	0.814

A definitive trend can be seen from table 3 that by increasing the size of the bag of words, results were greatly improved hence a better result classification can be observed.

c) Varying The Threshold: Next the variation was done for the threshold to extract the SIFT features from an image, while keeping the Vocabulary size and number of images constant as 500 and 10 respectively. Three values (0,5,10) of threshold were used for this computation. Table 4 shows the AUC obtained.

Table 4: Results for varying the threshold value

	0	5	10
Bicycle	0.649	0.720	0.667
Bus	0.498	0.540	0.518
Car	0.728	0.646	0.502
Cat	0.495	0.609	0.568
Cow	0.701	0.644	0.663
Dog	0.496	0.647	0.600
Horse	0.452	0.568	0.598
Motorbike	0.542	0.568	0.597
Person	0.494	0.683	0.518
Sheep	0.607	0.600	0.495

When we increase the value of the threshold, we are actually rejecting peaks with low absolute values. The number of features extracted will be less for high values of threshold but features will be of good quality. Thus choosing high threshold implies that the number of features extracted will be less and more confined and we may not be able to classify the object classes efficiently. From the results, shown in table 4 it is obvious that different classes have good results for threshold value in the range 0 to 5. This is because for threshold value of 0, we are accepting all the features and results are not good for difficult classes like dog, person etc. so in order to have better results for difficult classes features with fairly good quality must be accepted. This is the reason why difficult classes have shown good results with the threshold value of 5, in which we have accepted features of considerably good quality.

3.3 Dense SIFT

Next, the experiment was done with the dense SIFT by varying the size or scale of frames used to extract features from the image. As in Dense SIFT, we scan the whole image and features are extracted uniformly from each part (Frame or window) of the image, so by changing the scale, we effectively change the number of descriptors for an image. In this experiment, the vocabulary size was kept fixed at 500 and number of images used to build vocabulary had a constant value of 10. The scale of the frame had values of 1.5, 5 and 10. The results are shown in table 5.

Table 5: Results for varying scale in dense SIFT

	1.5	5	10
Bicycle	0.785	0.812	0.726
Bus	0.718	0.599	0.645
Car	0.756	0.736	0.708
Cat	0.689	0.667	0.661
Cow	0.841	0.801	0.783
Dog	0.665	0.524	0.501
Horse	0.638	0.498	0.498
Motorbike	0.583	0.595	0.595

Person	0.617	0.559	0.583
Sheep	0.695	0.662	0.629

Results came out quite good for a scale value of 1.5. Generally, as we increase the value of the scale, most of the classes don't have the good results as compared to low scale values. It is observed that Dense SIFT gives better results than Sparse SIFT but for some classes trend is opposite in which Sparse SIFT is better. DENSE SIFT has particularly given better result for the difficult class 'person' in which we get high value for AUC as compared to optimized sparse SIFT parameters. But for another difficult class, 'horse' results are not good for dense SIFT and sparse SIFT gives better results. So, it is difficult to conclude that dense SIFT is better than sparse SIFT, but still dense SIFT has an edge over sparse sift as results are quite good for some of the classes and average classification value is better for dense SIFT.

3.4 Covariance

Feature space is built with the 9 values from the eigenvectors of the covariance matrix. The results of this technique were not good. It was expected that objects may have covariance with their respective background types. But the results were against this as most of the classes have very low AUC. The results as shown in table 6 are highly unexpected especially for cow which usually has high AUC with other methods. The main reason could be the varying and diverse nature of backgrounds for each object class.

Table 6: Results for covariance

Bicycle	0.467
Bus	0.653
Car	0.559
Cat	0.628
Cow	0.496
Dog	0.552
Horse	0.598
Motorbike	0.605
Person	0.438
Sheep	0.583
Average	0.5579

4. Classifiers

Performance of different classifiers was also tested in order to achieve better results.

4.1 Varying Classifier Type

Here, the main comparison is done between KNN classifier and Support Vector Machine SVM classifier. 10 images were used to build a standard vocabulary of 500 words. The value of the threshold is kept constant at 0. Next, different classifiers are selected from the 'PRtool box' and their performance is tested with the same dataset.

Table 7: Results for varying classifier types

	KNN-1	SVM-P1	SVM-R1
Bicycle	0.412	0.625	0.627
Bus	0.489	0.365	0.381
Car	0.312	0.685	0.697
Cat	0.624	0.649	0.632

Cow	0.739	0.673	0.675
Dog	0.592	0.493	0.484
Horse	0.457	0.394	0.491
Motorbike	0.513	0.515	0.521
Person	0.516	0.628	0.623
Sheep	0.565	0.725	0.729

The results of the experiment can be seen from table 7. It was observed that SVM classifier clearly performs better as compared to KNN classifier.

4.2 Combining Different Classifiers

In this we combined different classifiers together. The classifiers combined were SVM, KNN and LDC.

Table 8: Results for combining different classifiers

Bicycle	0.322
Bus	0.522
Car	0.306
Cat	0.664
Cow	0.710
Dog	0.601
Horse	0.394
Motorbike	0.467
Person	0.532
Sheep	0.618

From table 8, it is observed that combining different classifier's didn't produced good results. The performance of the individual classifiers was better compared to combined ones.

5. Combining Features

In sparse SIFT, we only take strong features and in Dense SIFT, we take features from a window scanned over the whole image. This thing brought an idea that the performance of both these features should be checked together. So for this purpose the sparse and dense SIFT features were combined.

Table 9: Results for combined Sparse and Dense SIFT

Bicycle	0.792
Bus	0.568
Car	0.765
Cat	0.753
Cow	0.849
Dog	0.652
Horse	0.514
Motorbike	0.634
Person	0.552
Sheep	0.714

From table 9 it can be seen that the performance is quite good as compared to the individual case. The main reason for this can be that by combining sparse and dense SIFT; features from all over the image are taken along with the strong features. This might have helped the classifier to draw better decision boundaries, considering the image is now represented in a better way.

6. Optimization

Keeping track of all the experiments done in the previous sections and their optimized parameters, a final classifier is built. As both sparse and dense SIFT show competitive results and it was difficult to decide about the selection of one, final classification is done with both sparse and dense SIFT. For sparse SIFT, the value of the threshold is taken to be '0' and for dense SIFT, the value of the scale is taken to be '1.5' with the spacing step of '10'. This time the size of the vocabulary is taken as 500. Vocabulary size of 1000 and 1500 can also be used for further improvement of results. The results of the final classifier are shown in table 10.

Table 10: Results for final classification

	Sparse SIFT	Dense SIFT
Bicycle	0.792	0.817
Bus	0.667	0.725
Car	0.745	0.756
Cat	0.598	0.649
Cow	0.815	0.848
Dog	0.647	0.549
Horse	0.614	0.503
Motorbike	0.539	0.642
Person	0.683	0.540
Sheep	0.814	0.725

7. Conclusion

Object classification is a difficult and complex problem. There is no one technique which outclasses other techniques in all cases. Further, the performances of the classes are not uniform for any given technique. Different classes behave differently with the changed set of parameters. So, it is difficult to devise a single strategy which is equally good for all the classes.

In general, it is found that SIFT performs better than other techniques where dense SIFT has a slight edge over sparse SIFT. The number of images used to build vocabulary also affects the results. The greater the number of images better will be the results. But after a certain value, increase in number of images have a very small impact on result improvement but adds a huge computational cost. The size of the vocabulary is also a very important factor and results improve a lot with the increase in the vocabulary size. Bag of words is a smart technique and has proven quite good in object or scene classification techniques. Also it is observed that quality and quantity of the SIFT features also effects the performance of the classifier. The SVM classifier was found to be better than KNN classifier for this case. The quality (in terms of diversity and description for a particular class) of images used for the training and bag of words is also an important factor for classification performance.

References

- [1] M. Everingham, A. Zisserman, C. K. I. Williams, and L. Van Gool. The pascal visual object classes challenge 2006. PASCAL Network of Excellence on Pattern Analysis, statistical Modelling and Computational Learning.
- [2] R. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, and D. Tax, *PRTTools4, A Matlab Toolbox for pattern Recognition*, Delft University of Technology, February 2004.
- [3] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms,"
- [4] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," in *In Proc. 9th European Conf. on Computer Vision*, 2006, pp. 589-600.

Hammad Naeem received his degree of Master in Computer Vision and Robotics from Heriot Watt University-Edinburgh, University of Girona-Spain and University of Burgundy-France. He was on an Erasmus Mundus program which was funded by European Union. He holds Bachelors in Electrical Engineering from Air University, Islamabad, Pakistan. Currently he is serving as a Lecturer in Department of Electrical Engineering at HITEC University, Taxila, Pakistan. He has particular interests in Vision based robotics and image processing based intelligent machines.

Maria Minhas is currently doing her Master's in Electrical Engineering from University of Minnesota, Duluth, USA. She did her Bachelors in Electrical Engineering from HITEC University, Taxila, Pakistan. She has research interest in image processing and Bio-medical vision based solutions.

Jameel Ahmed received BSc. degree in Electronic Engineering from NED University of Engineering & Technology, Karachi, Pakistan in 1991. He did MSc. Electrical Engineering (Telecommunication) from NUST, Pakistan and PhD (Telecommunication Engineering) from Hamdard University Pakistan and Nanyang Technological University (NTU) Singapore. Subsequently, carried out Post Doctoral Fellowship from NTU Singapore. Jameel Ahmed is actively involved in teaching and research for the last twenty years. Presently, he is serving as professor and chairman, department of Electrical Engineering at HITEC University, Taxila Cantt-Islamabad, Pakistan. Dr Jameel is also heading the Signal Processing & Communication Systems (SPCS) research group of the department.

Cloud Computing: A Survey on its limitations and Potential Solutions

Mohammad Manzurul Islam¹, Sarwar Morshed² and Parijat Goswami³

¹Faculty of Engineering and Information Technology, University of Technology Sydney (UTS)
Sydney, NSW 2007, Australia

²Department of Computer Science and IT, Islamic University of Technology (IUT)
Gazipur, Dhaka, Bangladesh

³Central Operations Division, Robi Axiata Limited
Dhaka, Bangladesh

Abstract

Cloud computing has brought a new paradigm shift in technology industry and becoming increasingly popular day by day. The small to medium enterprises (SMEs) are now adopting cloud computing in much more higher rate than large enterprises. That raises a debate whether this cloud computing technology will penetrate throughout the IT industry or not. The SMEs are adopting cloud computing for the low cost implementation of total IT infrastructure and software system whereas the large enterprises are relying on their own infrastructure for data security, privacy and flexibility to access their own infrastructure. In this paper, we provide a survey about possible limitations of cloud computing that is delaying its penetration. We also identify the ongoing potential solutions that will help the enterprises to adopt cloud computing for their IT infrastructure and software systems.

Keywords: *Cloud Computing, Enterprise, SME, Limitations, Survey, Potential Solutions.*

1. Introduction

Cloud computing is a modern computing technology where software and hardware infrastructure of an enterprise can be placed over a network to access later in on demand basis via internet instead of having them locally within the enterprise. Cloud computing service provider holds the responsibility to manage and share all the hardware and software using virtualization among the clients, and the clients only pay for the subscribed services [1][12]. Cloud computing turned out to be useful for small to medium enterprises (SMEs) in order to have low implementation cost for their total IT infrastructure and software systems. For SMEs, the pay per user basis service license drastically reduces the cost of both hardware and software [15][17]. Despite of many benefits of cloud computing towards small to medium enterprises (SMEs), large organizations are still skeptical to adopt cloud-

computing services and still relying on their own infrastructure for data security, privacy, reliability and flexibility issues. Evidently, cloud computing comes with few potential limitations that is delaying its adoption in the IT industry [6].

This paper explores all the possible drawbacks of cloud computing and lists all the prospective solutions available till today. Thus SMEs and large enterprises will be able to take decision about adopting cloud computing and the readers will have a clear overview about the future penetration of cloud computing in the IT industry.

2. Adoption of cloud

Cloud computing is adopted mostly by the SMEs due to its numerous benefits. By utilizing the tremendous abilities of cloud computing, SMEs can deploy applications lot quicker and cheaper compared to the cost of setting up whole IT infrastructure and services by themselves. Moreover, cloud computing does not limit on software license – small numbers of users get the same benefit as larger number of users [1][13].

An enterprise must check several things in order to move into the clouds. Firstly, the customer must technically comply with the existing cloud system. Secondly, moving the data to the cloud should not violate any security law of the nation or break the customer data privacy policy. Thirdly, the internal must allow executing the workloads on the cloud environment. Finally, companies must prepare their business process in a way that on demand cloud products can be acquired whenever required [9].

However, large enterprises have enormous private resources to keep their data private and secure. Furthermore, these large companies want to control their own total IT infrastructure by themselves to

provide reliable service to their customers, rather than risking their reputation to a third party cloud service provider. Therefore, large enterprises are not adopting cloud computing [1][16].

3. Limitations of cloud computing

Enterprises observe different obstacles when they move their IT infrastructure into the clouds. SMEs can sacrifice the sufferings of these obstacles to some extent, since adopting cloud would be a cheaper solution compared to the cost of running an individual IT infrastructure. A perfect trade-off between costs and benefits can help SMEs to make proper judgment of adopting cloud computing [5]. In this section, we analyze all the potential problems that delay the adoption cloud computing for some of the SMEs and the large enterprises.

3.1 Privacy and Security of data

Privacy and security are the two main concerns related with the adoption of cloud computing. As the resources are distributed among different cloud clients, privacy and security of data faces severe threat. An eavesdropper or unethical client can become a potential threat towards the normal users.

Moreover, clients have to completely depend on cloud service provider companies for their data security. The data administrator of the cloud provider can easily manipulate the sensitive data. For smaller organizations with limited resources, trusting a cloud provider may be safer than keeping them on premises.

Nevertheless, for larger organizations, cloud providers cannot provide sufficient protection against safeguarding of sensitive data of the customer, trade policy, business strategy, price information or any other classified information [1][16]. For example, an online bookmarking organization encountered a serious data failure in 2009 that resulted into total loss of client bookmarks. Also in 2008, a cloud vendor named 'The Linkup' having more than 20,000 paying customers, have lost most of its customer's data after a tragic system crash. The Linkup went out of business after this catastrophic system failure. They blamed their problems on their storage partner company but could not do anything else to rectify customer's loss [18].

3.2 Vendor Lock-in and Interoperability

Cloud computing software and hardware platforms are vendor dependent. Since different vendors use different software and hardware architecture for their own system, migration of user data and service from one vendor to other is nearly impossible. A client, that uses two separate vendors for its IT services, cannot

integrate different services between two vendors. This phenomenon is known as vendor lock-in.

Data portability, migration and vendor lock-in situation will increase with the rise of cloud computing. Many providers are now entering into the cloud market with their own software and hardware solutions. Furthermore, the clients have no control over the IT infrastructure and software services once all the services are in the cloud. The client then have to rely solely on the service providers, consequently they no longer have full control over their own IT [6].

All the major cloud providers have their proprietary data storage. For example, Google uses BigTable, Facebook uses Cassandra and Amazon uses Dynamo. There is no common interface to access these databases. Migration of contact data from Salesforce to Gmail and vice versa is not possible due to the absence of common interface and database systems. This leads cloud computing into vendor lock-in state and data cannot be migrated in this situation [6].

3.3 Service availability

Every single organization, whether SMEs or large, want 100% availability of the services they subscribed to. SMEs might accept minor unavailability of service when they make trade-off with the cost. Service like VOIP needs constant uptime. A small fragment of missing data may result in strong customer dissatisfaction. Even organization like Google or Amazon sometimes might face service outage. In contrast, if a cloud customer cannot access his services in the cloud, he has nothing to do except waiting for the service to be made online by the service provider. Aside from technical faults, a cloud provider can also face service unavailability due to going out of business or being the target of regulatory actions [2]. This in turn will have huge effect on all of its clients.

3.4 Absence of proper Service Level

Agreement (SLA)

Lack of well-defined service level agreement (SLA) is another burning issue of cloud computing. Currently there is no standard SLA for guiding the services provided by a vendor and each cloud vendor has their own defined SLA. Absence of proper SLA is the barrier to data migration among different vendors. It also keeps the client into dark about the services they could expect from any cloud vendor. Furthermore, lack of interoperability keeps the customer locked into a single cloud vendor [8].

There are different efforts of preparing standard SLA for cloud computing. Some are led by the vendors, while others by different standard body. As a result, there is no single point of control for implementing a universal SLA. Customer input can help

standardization to meet satisfactory level of service because a customer can note down its requirement towards a vendor. As cloud computing is evolving every day, an SLA prepared today can be obsolete tomorrow due to the rapid changing behavior of the clouds [8].

As the cloud market observed perfect market competition, pricing pressure results into decrease of guaranteed level of performance, service uptime and vendor responsiveness. For example, a well-known cloud provider can offer service up-time of 99.99% along with 10 minutes of downtime per year with a 10% discount on service charge on the month that fails to meet this requirement. In this way, a cloud vendor is offering certain discount knowing that their infrastructure will not be able to meet that up-time requirement in exchange for the benefit of claiming the level of service reliability. If a customer really needs the 99.99% up time, then this discount will not help him from facing severe loss of revenue [4][6].

3.5 Performance Instability

Cloud computing observes severe performance instability during high load. The expected behavior cannot be predicted as the cloud resources are shared among different users with variable process load. Some researches in Australia conducted different stress test on Amazon, Microsoft and Google to demonstrate the variations in service availability and performance due to variable load. They have measured the consequences of sudden demand of 2,000 simultaneous users and found that the response time varies by a factor of 20 at different points of the day [6].

3.6 Latency on network

The internet data transfer rate is relatively lower than the Ethernet data rate. Nowadays, fiber optic cable is replacing the local networks and some parts of the wide area network. Fiber optic uses light to send data; hence, data travels at the speed of light. But the internet infrastructure cannot handle such high rate of data. Again, the applications use large volume of data. Transmission bottleneck is observed as the internet medium is transmitting large volume of data. For example, computer scientists in University of California, Berkeley, have measured the cost of sending 10 terabytes of data from Bay Area to Amazon in Seattle. On an average bandwidth internet link, it takes almost 45 days to transmit this large volume of data with a cost of USD \$1,000 as the network transfer fee. On the other hand, shipping ten 1 terabyte disks via any standard shipping service takes 1 day and costs USD \$400 only. Therefore, cloud computing is not suitable for sending large volume of data [2][6].

3.7 Lack of scalable storage

Scaling the storage system as required is not possible in a cloud computing architecture. Relational databases (RDBMS) are widely used in different organizations. In the cloud, there is no common interface to access different RDBMS at the same time. Cloud vendors have their own solutions for handling databases. If a customer requires more space in database but the cloud provider is out of space on its existing database, it cannot be increased just by adding an extra hard drive. Furthermore, updating of database requires longer time in the cloud. As a result, cloud computing might be inconvenient for data manipulation centric organizations [2][6].

3.8 Reputation fate sharing

A single customer's illegal behavior can affect the reputation and fate of other customers using the same cloud network. In the clouds, all the clients of the cloud provider share the same resources. A malicious user can get banned from a specific service that results into service unavailability towards the rest of the users on that cloud due to banning of internet IP address. Additionally, cloud providers transfer the legal liability to the users for performing illegal activities through their network. In March 2009, FBI raided a data center located at Dallas because one company, whose services were hosted there, was investigated for some criminal activities. At that time, all services on that data center were shut down and the other innocent customers suffered long downtime, for which some of them even went out of business [14].

4. Potential solutions

The problems mentioned in the previous section of this article currently do not have any precise solution. Many researchers around the world are working to solve the above specified issues with cloud computing.

i. Privacy and security: Privacy and security issues are two major obstacles to adopt cloud computing. Complex encryption method of data can provide more security. Research is going on developing such encryption method where users can send their data in encrypted bits. The cloud provider can search and calculate the encrypted data but cannot see what the actual data is. Theoretically such method do exists. But it requires high bandwidth and processor cycle, incurring higher expenses. Using of hardware based security can be more effective in respect to bandwidth and processor speed. But it also incurs higher cost [5][10]. Again, hosting the cloud infrastructure in more trusted region like European Union can add extra privacy and security towards the customers [1][11].

ii. *Vendor lock-in problem*: Vendor lock-in problem can be solved by enforcing all the vendors to use a unified API. It can result in decreasing of profit for the cloud vendors, but on the other hand more customers will now be interested to adopt cloud computing [2].

iii. *Service level agreement*: Proper SLA is far away from the light. Because the cloud computing itself is not mature enough to stick with a single unified SLA at this moment. Lots of changes are going on every day. All the vendors and standardization bodies should agree on making a unified SLA and work step by step rather than making their own SLA and convincing it to be accepted by others [8].

iv. *Performance instability*: Proper scheduling technique can solve performance instability problem. Research is on going to fix this problem [2].

v. *Scalable database*: Researchers are working to solve the issue of scalable database. Creating a new storage system with unified interface for all the cloud providers may be a solution [2]. A scalable storage with an SQL-like API is under construction to solve this issue [3].

vi. *Hybrid solutions*: Apart from the solutions mentioned above, a hybrid solution can also be adopted. In this case, an enterprise will only move the part of its IT infrastructure which has no effect from the regular problems of cloud computing. The remaining parts of the IT infrastructure will belong to the enterprise itself [7].

5. Conclusions

This survey paper explored different limitations of cloud computing and presented the ongoing potential solutions towards those problems. The fundamental problems identified by most existing research are data privacy, security, vendor lock-in, interoperability, service availability, absence of unified SLA, performance instability, network bottleneck, lack of scalable storage and reputation fate sharing. Cloud computing is widely adopted by the SMEs for its low cost in spite of having such problems. On the other hand, large enterprises tend to rely on their own infrastructure rather than depending on cloud vendor. Since researchers are working to overcome the barriers of adopting cloud computing, soon most of the problems of cloud computing will be solved or the risk will be mitigated to an acceptable level. This literature indicates that there is much work to be done in developing the solutions. This is perhaps the most important concern of the future of cloud computing, as many enterprises might want to move their IT system into the clouds after a careful analysis.

References

1. Arinze, B. & Anandarajan, M. 2010, 'Factors that determine the adoption of cloud computing: a global perspective', *International Journal of Enterprise Information Systems*, IJEIS, vol. 6, no. 4, pp. 55-68.
2. Armbrust, M., Fox, A., Grieth, R., Joseph, A.D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I.O.N. & Zaharia, M. 2010, 'A view of cloud computing', *Communications of the ACM*, vol. 53, no. 4, pp. 50-8.
3. Armbrust, M., Lanham, N., Tu, S., Fox, A., Franklin, M.J. & Patterson, D.A. 2010, 'The case for PIQL: a performance insightful query language', *Proceedings of the 1st ACM symposium on Cloud computing*, Indianapolis, Indiana, USA.
4. Durkee, D. 2010, 'Why cloud computing will never be free', *Communications of the ACM*, vol. 53, no. 5, pp. 62-9.
5. Gatewood, B. 2009, 'Clouds on the information horizon: how to avoid the storm', *Information Management Journal*, vol. 43, no. 4, pp. 32-6.
6. Hofmann, P. & Woods, D. 2010, 'Cloud computing: the limits of public clouds for business applications', *Internet Computing*, IEEE, vol. 14, no. 6, pp. 90-3.
7. Khalidi, Y. 2011, 'Building a cloud computing platform for new possibilities', *IEEE Computer Society*, vol. 44, no. 3, pp. 29-34.
8. Ortiz, S. 2011, 'The problem with cloud-computing standardization', *IEEE Computer Society*, vol. 44, no. 7, pp. 13-6.
9. Ranjan, H. 2011, 'Cloud computing and EDA: Is cloud technology ready for verification (and is verification ready for cloud)?', *VLSI Design, Automation and Test (VLSI-DAT)*, 2011 International Symposium on, pp. 1-2.
10. Ryan, M.D. 2011, 'Cloud computing privacy concerns on our doorstep', *Communications of the ACM*, vol. 54, no. 1, pp. 36-8.
11. Bristow, R., Dodds, T., Northam, R., & Plugge, L. 2010, 'Cloud Computing and the Power to Choose', *EDUCAUSE Review*, 45(3), 14.
12. Cervone, F. H. 2010, 'An overview of virtual and cloud computing', *OCLC Systems & Services*, 26(3), 162-165. doi:10.1108/10650751011073607.
13. Cusumano, M. 2010, 'Technology Strategy and Management Cloud Computing and SaaS as New Computing Platforms', *Communications of the ACM*, 53(4), 27. doi:10.1145/1721654.1721667.
14. Fink, J. 2009, 'FBI agents raid Dallas computer business', <http://cbs11tv.com/local/Core.IP.Networks.2.974706.html>.
15. Greengard, S., & Kshetri, N. 2010, 'Cloud Computing and Developing Nations',

- Communications of the ACM, 53(5), 18.
doi:10.1145/1735223.1735232.
16. Lucky, R. 2009, 'Cloud Computing', IEEE Spectrum, 46(5), 27.
doi:10.1109/MSPEC.2009.4907382
 17. Mell, P., & Grance, T. 2010, 'The NIST Definition of Cloud Computing', Communications of the ACM, 53(6), 50.
 18. Wang, C., Ren, K., Lou, W., & Li, J. 2010, 'Toward Publicly Auditable Secure Cloud Data Storage Services', IEEE Network, 24(4), 5.
doi:10.1109/MNET.2010.5510914.

Mohammad Manzurul Islam received his M.Sc. degree in Information Technology with major in Internetworking from University of Technology Sydney (UTS) in 2013. He obtained his B.Sc. degree in Computer Science and Information Technology from Islamic University of Technology in 2005. He started his career as a Lecturer of Computer Science in Stamford University Bangladesh. At the same time, he was the Legal Main Contact, Curriculum Lead and Instructor of Cisco Local Academy. He also served in multinational telecommunication vendor (ZTE Corporation) and operator (Axiata Telecom Bangladesh). Currently he is working as the ICT Specialist of Abacus Services Limited, Bangladesh. His research interests include internetworking, network security and wireless sensor network.

Sarwar Morshed received his M.Sc. degree in Information Technology with major in Communication Engineering from Tampere University of Technology in 2009. He obtained his B.Sc. degree in Computer Science and Information Technology from Islamic University of Technology in 2005. He is currently Assistant Professor at Islamic University of Technology in Bangladesh. His research interests include wireless communication, internetworking and security in transmission.

Parijat Goswami received his B.Sc. degree in Electrical & Electronic Engineering from Bangladesh University of Engineering & Technology (BUET) in 2008. Currently he is working as the Manager, Central Operations at Robi Axiata Bangladesh Ltd (A multinational telecommunication company). His research interests include internetworking, artificial intelligence, wireless sensor network & telecommunication.

Intelligent Vehicle Recognition based on Wireless Sensor Network

Maha Mohamed Nabeel, Mahmoud FakhreEl-Dein¹, SherineAbd El-Kader
Computers and Systems Department

Electronics Research Institute,12622 El-Bohothst., Dokki, Giza, Egypt

Abstract

One of the main requirements of any intelligent transportation system is to be able to identify vehicles in the traffic. This paper presents an intelligent vehicle identification system used within a complete solution for a traffic monitoring system that uses a novel wireless sensor network architecture to monitor traffic. A novel wireless sensor network architecture to monitor traffic is proposed where a visual sensor node captures images of the traffic and sends them to the traffic control center for processing. Also, this paper compares between three main localization and recognition algorithms. To locate the vehicle logo in the traffic image a symmetry detection algorithm is used to detect the inherent symmetry in vehicle frontal images. A fine location of the logo is identified using three different methods in the region marked by the symmetry line. After locating the logo three feature sets are extracted and presented to the classifier to correctly identify the type of the vehicle. The results of the localization and recognition algorithms show the efficiency of the presented system in identifying vehicle types with a recognition rate over 90%.

Keywords: vehicle logo recognition, traffic control systems and wireless sensor network

1. Introduction

The massive increase in automobiles as one of the most important modes of transportation has facilitated human life but it introduces different types of problems as traffic congestion, parking problems and traffic accidents. Researchers are pursuing solutions through developing intelligent transportation systems to solve these problems. An intelligent transportation should be able to manage, monitor and direct users to a safer and more coordinated transportation. Primarily, an intelligent transportation system has to correctly identify, recognize and track vehicles in real-time [1]. This paper presents a solution to monitor traffic using Wireless Sensor Networks (WSN).

Wireless Sensor Network (WSN) is a network of small Sensors Nodes (SN) that could communicate between themselves and/or base location using a wireless communication system. A sensor node usually consists of a sensor unit, a computing and a communication unit.

Currently, researchers are increasingly using WSN due to its variety in function flexibility in development and numerous potential applications [2].

The architecture of a typical WSN contains a large number of SN and one or more Access Points (AP). The access points have higher computational power, enhanced communication capability and larger power source. In general, sensor nodes are small nodes that measure environmental conditions, deployed with a spatial density and at a high sampling rate that is defined by the application. The sensor nodes process the raw sensor data to extract useful information which is transmitted to the access point using the communication system. The access points could process the data collected from the sensors in the network to extract more relevant information then sends the data collected to the traffic operators or the traffic control system through the internet [3].

In traffic surveillance application, as shown in figure 1, sensors are positioned at predetermined locations on the pavement; their job is to gather the magnetic signature of vehicles travelling over the sensor nodes. The nodes use vehicle detection algorithms to process these signals. The detection events are then generated and transmitted to the base station from the entire network. Based on this collected information, the base station computes the vehicle count, occupancy and speed of the traffic flow and in turn transmits it to the control center. The traffic management center (TMC) uses this real-time information for road monitoring and traffic signals control [3].

This paper presents a complete system to monitor traffic and gather vehicle information using wireless sensor networks. The proposed system introduces the main architecture of the traffic control system. It provides a solution to traffic monitoring and vehicle identification using non intrusive sensor network. The vehicle identification system is presented and explained in details on this research as one of the main modules in traffic monitoring and surveillance system. Our proposed vehicle identification system is divided into logo localization and recognition.

¹ Crosspointing author – Mahmoud Fakhre El deen

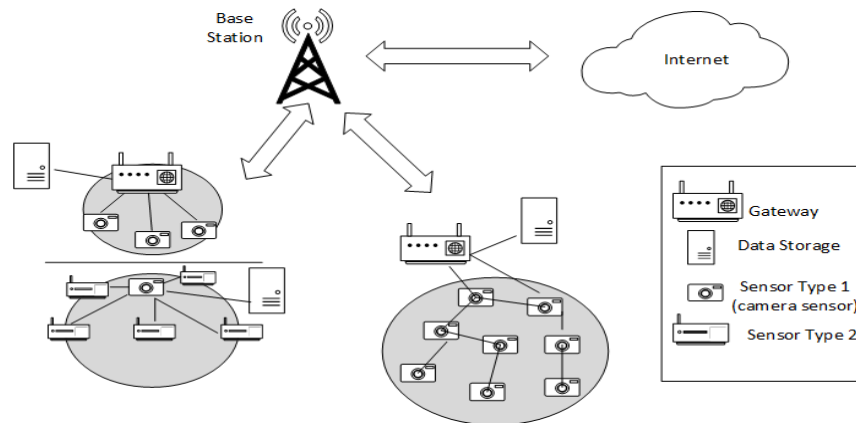


Figure 1: Typical wireless sensor network architecture

Vehicle logo localization locates the logo image in the input image and detects the fine location of the logo [4,5]. Firstly a course location of the vehicle logo is extracted from the input image by detecting bilateral symmetry of the vehicle front image. Next, the system tries to remove the background or vehicle grille using various textural analysis and edge detection methods. Based on the analysis of the previous step the fine location of the logo is detected the output of the vehicle logo localization is a fine location of the logo image [6]. The recognition of vehicle type from the logo image is divided into feature extraction and classification.

The rest of this paper is organized as follows; section 2 describes the different methods which used in traffic control using WSN and its challenges. Section 3 provides an overview about Bilateral symmetry detection and Section 4 provides a description of the proposed system and focuses on the vehicle detection and recognition strategies, focusing on the details of the algorithm's implementation. A set of experiments was used to evaluate the method's performance. The analysis of the results, achieved under various test conditions, is presented in section 5. Finally, section 6 presents the conclusions drawn from the work developed and highlights possible directions for future work.

2. Relatedworks

In this paper, a vehicle logo recognition algorithm for the wireless sensor traffic monitoring system is presented. In the next section, current research in traffic control using wireless sensor network is presented then various vehicle logo recognition algorithms is discussed.

2.1 Wireless Sensor Networks

Different architectures of WSNs in intelligent transportation system are demonstrated in [7]. The survey explains the difference between main architecture and presents a summary of traditional vehicle detection and surveillance technologies. One of the most notable systems is Sensys networks [8] Wireless Vehicle Detection System (WVDS). WVDS sensors are placed in the middle of a lane and transmit information to a repeater which is located near the traffic control cabinet. To increase the lifetime of the sensor and network an efficient communication protocol [9] is used. Traffic center gathers information from the access points to process it and analyze traffic real-time.

Using wireless sensor networks for traffic monitoring is investigated in [1] to provide data for Intelligent Transport System (ITS). They presented a case study in the freight village of Turin to evaluate the performance of a vehicle detection system based on magnetic sensors. WSN-based framework for traffic information collection according to the user requests is presented in [10]. Two key processes associated with the traffic information acquisition and delivery were specified using three types of sensor nodes. They used a two-layer hierarchical communication channel to establish communication among these sensors. A user-customizable data-centric routing scheme is proposed for data delivery, in which an efficient routing method is adopted for the routing decision-making. A real time intelligent transportation system capable of providing prompt information was presented in [11] based on acoustic arrays and powered by effective post-processing. An adaptive traffic control system using WSN is presented [12] based on controlling traffic on single and multiple intersections. The first part of this algorithm is a traffic sensor node installed on the roadside that collects periodically the traffic data from vehicles. The second part is a communication system algorithm that finds and

controls the communications between the sensor and the base station.

2.2 Vehicle Logo Localization

The vehicle logo localization is the first step in any logo recognition algorithms. The process is divided to coarse localization of the logo location then followed by a fine detection of the logo region. The coarse location of the logo is detected using prior information then a background and grille removal algorithm is used to finally detect the fine location of the logo. There are two main methods usually used in coarse location of the logo. The first is relative to the license plate location and the second is relative to bilateral symmetry line of the car. Various texture and morphological methods are used in background removal and fine location of the logo. The next section describes each method in more details.

2.2.1 Relative to the License Plate

Based on the prior knowledge that the logo usually located over the license plate, the course location of the logo is detected relative to the location of the license plate. Figure 2 shows a frontal image of a vehicle and the relative position of the logo to the license plate [3,13,14]. The approximate location of the logo is identified by an equation similar to Equation (1) .

$$\begin{cases} X_1 = X_{left} \\ X_2 = X_{right} \\ Y_1 = Y_{up} - N * width \\ Y_2 = Y_{up} \end{cases} \quad (1)$$

where X_{left} and X_{right} are the left and right position of the license plate, width is the width of the license plate and Y_{up} is the upper boundary of vehicle license .



Figure 2: The relative position of Vehicle logo to the License plate

Liu [4] presented this solution in his system which starts by locating the location of the license plate. After applying Equation 1 a coarse location of the vehicle logo is detected.

A texture recognition algorithm is used to detect the type of the grille of the vehicle. Liu categorized the grille to be horizontal, vertical or both. The next step is to remove this background based on which type of grille detected in the previous step. The final step uses morphological operators and vertical and horizontal projection to detect the fine location of the logo.

However, the aforementioned for License Plate Recognition (LPR) based approaches face the general problems caused by the viewpoint variation and/or non-symmetric frontal license plate as shown in figure 3. In essence, a universal solution to license plate recognition is still not available, it is highly dependent on prior knowledge about aspect ratio, background/character color and symbol types of the license plate, thus license plate recognition, in general, does not generate satisfactory results in countries or regions where vehicles have diverse license plates which might be issued by various authorities. Therefore, vehicle logo localization will eventually fail due to the incorrect LPR result [6].



Figure 3: Example of non-symmetric frontal license plate

2.2.2 Symmetry Detection

An easily recognizable feature of the front view of the vehicle is the air-intake grille which is mounted on most vehicle's front. It usually has a regular shape grid or horizontal/vertical bars which is easy to detect, co-occurring with vehicle logos and shares the same symmetric axis with the vehicle. Therefore, by detecting the symmetry line of the air-intake grille the potential position of the vehicle logo can be efficiently determined [6] .

This method is used in some systems as used by Zhou [6] who presented a system that relies on the feature-match based symmetry detection algorithm [15] for detecting bilateral symmetry in the image. Then used Hough transform to detect vertical and horizontal grille lines from the phase congruency edge image. The final location of the logo is detected using a sliding window inside the grille region.



Figure 4: The Result of Vehicle Logo Localization Phase

This method is more robust than localization relative to the license plate location but it has lower recognition rate. In this research we adopt this method and try to improve it using various background removal and fine localization of the system. The experiments implemented provided in section 5 an analysis of the achieved improvements.

2.3 VehicleLogo Recognition

The second step after localization of the vehicle logo is classifying the vehicle for a specific brand (Toyota, Ford ... etc.). This step includes extracting features from the logo image and comparing it to the logo against a database of known brands, and using the calculated similarities to decide the most likely manufacturer. Feature extraction is one of the key operations in any classification system. In vehicle logo recognition most systems either use moments or Feature Matching Schemes. This section illustrate how different systems used these feature extraction to identify vehicle types .

Moments are a certain weighted measure of image intensity which could be scaled, translation and orientation invariant. There is a different type of moment like Zernike moment, Hu moments and Legendre moment used in various recognition problems including gesture recognition and 3D object recognition [2]. Hu moments define seven

invariant moments to scale, orientation and translation. Unfortunately they have data redundancy problem because Hu moments are not orthogonal moments. Dai [2] used Tchebichef moments invariant as features for vehicle logo recognition. He compared Tchebichef moment with Zernike moment and Hu moment invariants.

Some promising approaches to detect foreground objects from images have been published: Scale Invariant Feature Transform (SIFT) [16, 17] that introduced in [18], Speeded Up Robust Features (SURF) [19]. Both methods detect interest points called features and propose a method of creating an invariant descriptor for these features. The authors of SURF claim to be a superior to SIFT in term of runtime execution while it is still providing good results with regards to feature point quality [16]. The created descriptor is used as a vector, uniquely identifying the found interest points. It has to be distinctive and robust for various scale-space deformations. Vectors can be used for matching indetecting interest points even under a variety of disturbing conditions like scale changes, rotation, changes in illumination and viewpoints or image noise. The invariance is the most important ability of these key-point detectors.

Different classification method was evaluated in vehicle recognition systems. In [20], a comparison between various classifiers and classifier fusion methods is presented. They compared between Support Vector Machine (SVM), Back propagation Neural Network and K-nearest neighborhood classifier as they present the most used classifiers. Others used Euclidean distance measure to match key point of the input logo to the key points of the database [17]. In the next section we will present a necessary background information about the main algorithms used in this research. The bilateral symmetry algorithm and SIFT feature.

3. Bilateral Symmetry detection

The first step in the system is bilateral symmetry to detect the symmetry in the vehicle front image. This process was first adopted by Zhou [6] based on symmetry detection presented by [15]. The algorithm is divided into three main steps : SIFT descriptor computation, symmetry descriptors detector, symmetry line estimation. The first step in the algorithm is computing the SIFT features an extracting the descriptors for keypoints (see next section3 for details of the algorithm).The next step starts by finding similar or mirror keypoint based on the values of the SIFT descriptor. The process removes all key points that do not have mirror feature in the image. This step also includes grouping keypoints to a set of mirrored keypoints. The final step starts by constructing small symmetry lines between each two pairs of mirrored points and then tries to detect symmetry line by generating a Hough transform of all the symmetry line from all mirrored keypoints. More than one

symmetry line could be detected and so a confidence value is associated with the symmetry line based on the strength and number of the mirrored keypoints.

3.1 Scale Invariant Feature Transform (SIFT)

Matching features across different images in a common problem in computer vision. When all images are similar in nature (same scale, orientation, etc.) simple corner detectors can work. But when you have images of different scales and rotations, you need to use the Scale Invariant Feature Transform.

Scale Invariant Feature Transform (SIFT) is an image descriptor for image-based matching developed by David Lowe [21]. This descriptor as well as related image descriptors is used for a large number of purposes in computer vision related to point matching between different views of a 3D scene and view-based object recognition. The SIFT descriptor is invariant to translations, rotations and scaling transformations in the image domain and robust to moderate perspective transformations and illumination variations. The approach is efficient on feature extraction and has the ability to identify large numbers of features.

SIFT is implemented as the following stages: Creating the Difference of Gaussian Pyramid, Extreme Detection, Keypoints Elimination, Orientation Assignment and Descriptor Computation. Firstly, a Gaussian scale space $D(x,y, \sigma)$ is constructed from difference of Gaussian images from the convolution of input image $I(x,y)$ with Gaussian filter with different widths or (sigmas σ). Next the goal of extrema detection is to find the extrema points in the DOG pyramid. To detect the local maxima and minima of $D(x,y, \sigma)$, each point is compared with the pixels of all its 26 neighbors. If this value is the minimum or maximum this point is an extrema.

Key point elimination attempts to eliminate some points from the candidate list of key points by finding those that have low contrast or are poorly localized on an edge [21]. Next, the orientation assignment step, assigns consistent orientation to each key point based on the local image properties. The key point descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation. SIFT is orientation based on the orientation histogram which is formed from the gradient orientations at all sample points within a circular window around the key point. Each sample added to the histogram is weighted by its gradient magnitude and by a Gaussian-weighted circular window with a three times that of the scale of the key point. The orientation histogram has

36 bins covering the 360 degree range of orientations. The final step, descriptor computation is computed for the local image region that is as distinctive as possible at each candidate key point. The image gradient magnitudes and orientations are sampled around the key point location.

4. System Description

This paper presents a wireless sensor network architecture to monitor traffic using a vision based algorithm. Figure 5 shows the architecture of the proposed system. The figure shows that the traffic is monitored using two types of sensor nodes a small sensor nodes mounted on each streetlight and more powerful access points mounted on the traffic light of each street. The network depends on non intrusive sensor nodes that gather traffic and vehicle data to reduce the cost of installation on current traffic systems. The sensor will basically build based on acoustic sensor implementing a vehicle detection algorithm similar to [11]. The sensor nodes transmit information from the sensors to the access points which contain a camera sensor, higher communication unit and more powerful computing unit. The data are transmitted using LLEAP routing protocol [22]. The access points detect the activity in the street by utilizing information from the neighborhood sensors. The access points transmit the data to the traffic control unit when activity is detected in the street. The traffic center relieves street and vehicle images and sensor data gathered from the whole network to extract full traffic monitoring information and statistics. The first step our automatic monitor system is vehicle logo recognition algorithm to detect the vehicles in camera feeds from the data in the traffic monitoring center.

One of the main identifying features of the vehicle is the vehicle logo. It contains the information about the vehicle model and manufacturer which could not be replaced easily. This provides an important basis for identifying and tracking vehicles accurately. Vehicle logo recognition is mostly divided into two main steps logo localization and vehicle logo recognition. Although vehicle logo localization seems an easy problem however the inherit particularity of vehicle logos makes accurate localization a more challenging problem.

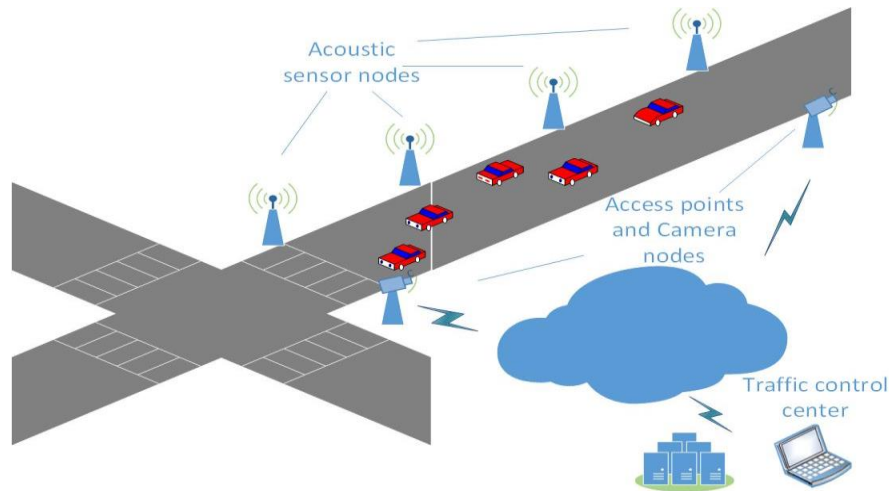


Figure 5: Traffic monitoring system

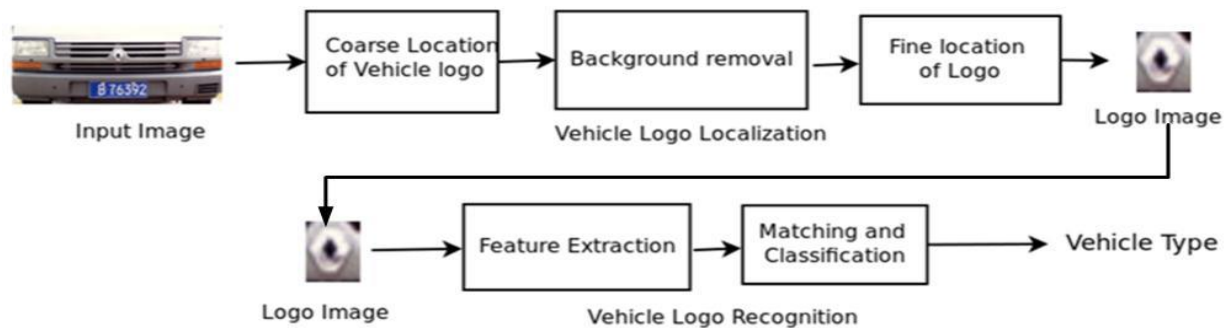


Figure 6: Basic Block diagram for any Vehicle Recognition System.

The next section presents in details the vehicle logo recognition system that identifies vehicles captured by a camera sensor in our traffic monitoring system. Figure 6 illustrates the main steps in any vehicle recognition system; the first step is vehicle logo localization which locates the logo image in the input image. The second step, vehicle logo recognition, classifies the a logo image to one of known vehicle types. In most systems vehicle logo localizations is divided into coarse location of logo, background removal and fine location of logo [4,5]. Firstly the coarse location of the vehicle logo is extracted from the input image. Next, the system tries to remove the background or vehicle grille using various textural analysis and edge detection methods. Based on the analysis of the previous step the fine location of the logo is detected the output of the vehicle logo localization is a fine location of the logo image. The recognition of vehicle type from the logo image is divided into feature extraction and classification.

4.1 Localization

The system is divided into logo localization and logo classification . Figure 7 shows the block diagram of the

localization algorithm. The first step in the localization algorithm is the bilateral symmetry detection which detects SIFT features and use it to detect the symmetry. The bilateral symmetry algorithm is similar to implemented by Zhou [6]. The algorithm uses SIFT features to detect major features in the image. The next step focus on matching the mirrored features together and removing all non mirrored. The mirrored featured are matched and a rough symmetry line for each set is drawn into the images. Hough transform is used to detect the symmetry line with the highest confidence in the image.

Figure 8 shows different stages of the bilateral symmetry algorithm. The first figure shows the SIFT key features extracted from the image. The next picture shows only the matched mirrored features after removal of all non mirrored features (the green lines in the image). Figure 8 (a) shows the symmetry axes of each two matched features. This image is used as input for the Hough transform which produce output similar to figure 8 (b) . This image shows the main matched mirror features along the symmetry line with the highest confidence. The next step after detecting the symmetry axis of the image is locating the logo in the input image. The system is based on the assumption that the logo

is directly over the grille and that we are detecting the logo from the front of the image. The easiest method of locating the logo from the input image is to crop the input image based on a margin of the symmetry line. The result of this cropping is illustrated in the following figure9.

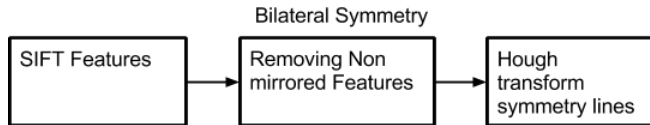
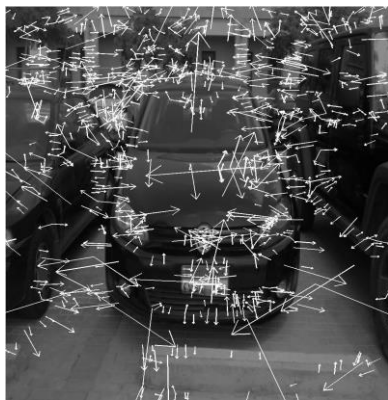


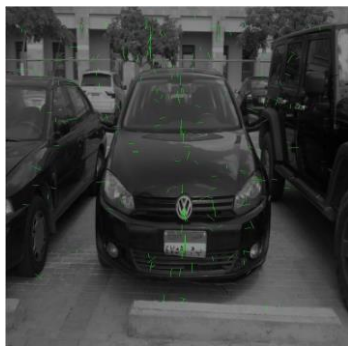
Figure 7: Block diagram of the localization algorithms



(a) SIFT features



(b) Mirror Features



(c) Symmetry Points



(d) Final stage

Figure 8: Stages of Bilateral symmetry detection



Figure 9: Final symmetry axis detected

4.1.1 Hough Transform based algorithm

A variation of the method presented in [6] was implemented where Hough transform is used to locate the region of the grille and then a horizontal and vertical filter applied in this region of interest to fine detect the logo location. Figure 10 shows sample of the output of this system on the dataset, these result confuse the background grid and buildings with the requested car grille location and move the solution location from beside the symmetry line to a wrong location. A different approach was used to solve the problem in [6] algorithm.

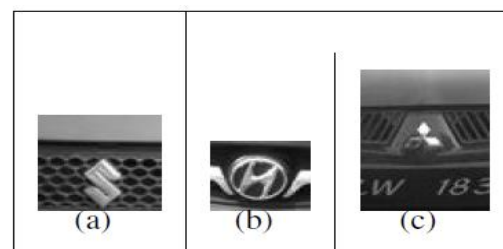


Figure 10: Result of Hough transform to detect the grille and then gradient to detect logo location

4.1.2 Texture analysis algorithm

The second method that was implemented extracted the textural features of the grille in location beside the symmetry line. In this experiment we used some Gabor filters [23] and

then use a k-means clustering algorithm to divide output of the 2D Gabor filter into different textural regions. Each region is then examined to check if it contains the grille. This examination is applied using sliding window and a horizontal and vertical line detection mask. Figure 11 shows sample of the result extracted using this procedure.

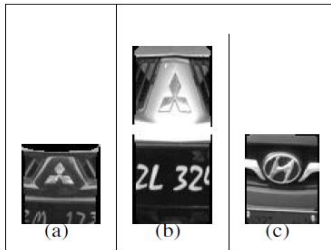


Figure 11: Example of the output of the Gabor filter texture analysis

4.1.3 Gradient algorithm

The third algorithm used to fine locate the logo image is a gradient based searching of the grille. After detecting the course location of the logo from the symmetry line. The system applies two gradient filters one to detect horizontal line and the second to detect vertical lines. After the filters are applied the system undergoes some morphological operations to fill the gaps and create a better mask for the air grille in the image. The final step is searching for the location of horizontal and vertical lines in the mask images. This final step produce and image with the location of the logo and the grille clear from the background. Figure 12 shows samples of the result extracted using this procedure.

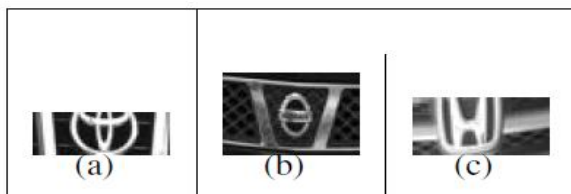


Figure 12: Example of the output of the gradient algorithm

4.2 Recognition Algorithms

After the vehicle logo is located in the input image and a final logo image is extracted by the logo localization algorithm. This is the image that is used as an input for the next phase of the feature extraction and classification. A composite set of features is used to generate a three main feature vector. The following list describes each feature vector in details.

4.2.1 Scale Invariant Feature Transform

SIFT features is used to extract the key points from the logo image, the SIFT has an advantage to be robust to changes and will generate similar features for small and large

images. This is helpful as the logo image may contain some other background. The SIFT feature are then concatenated into single features. To solve the problem of unbalanced features from different Images and types we only use the largest N^2 features in all input images and if the number of key descriptors is less than N we add a dummy zero value feature to the feature vector.

4.2.2 Gradient Operators

These features use global gradient with Gaussian mask: The gradient of the image is measure of the intensity and direction of the image. The systems starts by applying a simple gradient operator then computes the direction in 8 different sublayers finally each sublayer is convolved using a gaussian mask to generate the final feature set.

4.2.3 Wavelet

Our proposed solution uses wavelet transform to extract key features in the logo image. *2D wavelet transform*: A transform that captures both frequency and location information.

After computing the features, the symbol is introduced to the classifier as a feature vector. The system uses Support Vector Machine (SVM) classifier with Linear kernel [24]. An OVO classifier structure (object versus object) is used to handle the multi - class problem.

5. Results

This section presents the experiments conducted on the system and explain the obtained results. The dataset used in testing system was gathered as a dataset for license plate [25]. It consist of 708 frontal car images and 710 rear car from 8 different car manufacturers. We only used the frontal car images because the main constraints of detecting the logo on top of the grille. After the first few experiments we found that the dataset has images of cars without an identifying logo. A manual removal of images with no logo or when the logo is too small with respect to the car in the image. We also collected a set of 50 images to increase the dataset and compensate for removing images. Finally, we had 560 images that constitute our dataset used in localization and recognition.

In the experiments we conducted on the localization algorithm the test set implies that whole 560 images. In the recognition phase, we divided the data set into 70% training and 30% testing dataset.

Figure 13 shows a result of the bilateral symmetry block with the line of symmetry and the main symmetry featured marked on the input image. In all the images the number of images that the logo is located in the image is 93.82%. The

² Empirically determined to be by 74 feature

figures below show examples of a correct and an error locating of the symmetry line and cropping of the logo. Most errors in detecting the symmetry lines are due to confusion between the car symmetry features and dominate symmetrical background features. Other errors include two short or a too long symmetrical line in the front of the car.



Figure 13: Sample of Correct localization

Figure 14 shows the results of localization using the different algorithms. The table shows the best result achieved by the system is 93.82% using Simple cropping.

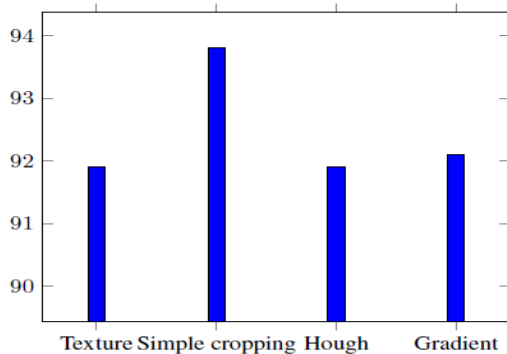


Figure 14: Final Localization Results

In Recognition phase, the dataset is divided into training, and testing is set to 70, 30 % respectively. A linear Support Vector Machine SVM classifier is used as a classifier for the

system. The results presented for the SIFT Features are similar to the 91% in [17]. Figure 15 shows the result of each feature extraction method. It is clear from the table that gradient provide better results than raw wavelet and SIFT features. This is natural as the gradients are powerful sets of feature and capture the local difference with more accuracy than wavelet features.

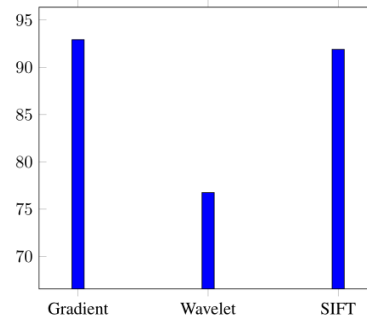


Figure 15: Final Recognition Results

6. Conclusion and Future Work

This paper presented a vehicle identification used in a traffic monitoring system using wireless sensor network. The architecture of the system consists of a wireless sensor network monitoring the traffic using various types of non-intrusive sensors. Acoustic and basic sensor nodes are mounted near each streetlight. These nodes has minimum processing power that sends information to larger access nodes. The access nodes are installed on the traffic lights and will contain a camera sensor, larger processing unit and wireless access point to connect to the traffic control center. The access node process the information from sensor nodes and camera sensor to detect vehicles. The detection of vehicles triggers the transmission to the traffic monitoring center. This paper focus only on vehicle localization and recognition algorithms that are used in our traffic monitor system.

The vehicle logo localization algorithm based on the symmetry of the frontal image of the vehicle as well as the air-intake grille as the dominant features has been presented. The system uses SIFT key features to detect the mirror features along y-axis in the image. This symmetry line is line with the most number of mirrored features in the image. Three different logo localization algorithms are presented to fine locate the logo in the image. The first is just simple cropping with some empirically margins. The second uses gradient horizontal and vertical masking to detect horizontal and vertical lines in the grill. The third algorithm scans the input window and tries to detect vertical and horizontal lines using Hough transforms. A comparison of the three algorithms is presented in the result with final accuracy rate of 93.82% , 92.18%, and 91.97% respectively. The recognition of the logo is tested using SVM classifier with

two different features set, gradient and wavelet transform. The final recognition rate of the system is over 90% which is better than current vehicle logo recognition systems.

This paper represents a proposed architecture for a wireless sensor network to monitor traffic with the a detailed investigation of the vehicle recognition algorithms. The research focus on introducing vehicle logo localization and recognition algorithms to identify vehicle type in the traffic monitor procedure. Further work on the system will start by building the sensor and access point nodes. After building the nodes the next steps will include various simulations of power consumption and various vehicle detection algorithms using acoustic and magnetic sensors.

Further simulation of the wireless sensor network is needed to test the power consumption and effective distance between nodes. Also, further improvement to the current vision system could be included for testing other feature extraction methods and implement a more domain specific features. This will enable fast feature extraction that will improve execution time. Another improvement is to extend the system to deal with logo localization for a vehicle with irregular shaped air grille, back view of the vehicle instead of frontal view only

References

- [1] S. Dai, H. Huang, Z. Gao, K. Li, and S. Xiao, "Vehicle-logo recognition method based on tchebichef moment invariants and svm," in WRI World Congress on Software Engineering, vol. 3, (Xiamen, China), pp. 18–21, Ieee, 19-21 May 2009.
- [2] Sherine M. AbdelkaderBasma M. El-Basioni and Mahmoud A. Fakhreldin, "Smart home design using wireless sensor network and biometric technologies," International Journal of Application or Innovation in Engineering and Management (IJAIEEM), vol. 2, pp. 413– 429, March 2013.
- [3] M. Bottero, B. D. Chiara, and F. Deflorio, "Wireless sensor networks for traffic monitoring in a logistic centre," Transportation Research Part C, vol. 26, pp. 99–124, 2013.
- [4] Y. Liu and S. Li, "A vehicle-logo location approach based on edge detection and projection," Vehicular Electronics and Safety (ICVES), pp. 165–168, 10-12 July 2011.
- [5] W. Yunqiong, L. Zhifang, and X. Fei, "A fastcoarse-to-fine vehicle logo detection and recognition method," in ROBIO 2007. IEEE International Conference on Robotics and Biomimetics, (Sanya), pp. 691 – 696, 2007.
- [6] K. Zhou, K. M. Varadarajan, M. Vincze, and F. Liu, "Hybridization of appearance and symmetry for vehicle-logo localization," in 2012 15th International IEEE Conference on Intelligent Transportation Systems, (Hilton Anchorage, Anchorage, AK, USA), pp. 1396–1401, IEEE, 16 - 19 September 2012.
- [7] M. A. Kafi, Y. Challal, D. Djenouri, A. Bouabdallah, L. Khelladi, and N. Badache, "A study of Wireless Sensor Network Architectures and Projects for Traffic Light Monitoring," Procedia Computer Science, vol. 10, pp. 543– 552, jan 2012.
- [8] "<http://www.sensysnetworks.com/home.>," accessed 2013.
- [9] P. Ergen, S.C. Varaiya, "Pedamacs: power efficient and delay aware medium access protocol for sensor networks," IEEE Transactions on Mobile Computing, July 2006.
- [10] J. Zhou, S. Member, C. L. P. Chen, L. Chen, and W. Zhao, "A user-customizable urban traffic information collection method based on wireless sensor networks," IEEE Transaction on Intelligent Transportation Systems, pp. 1–10, In press 2013.
- [11] B. Barbagli, L. Bencini, I. Magrini, G. Manes, S. Marta, and A. Manes, "A real-time traffic monitoring based on wireless sensor network technologies," in 2011 7th International Wireless Communications and Mobile Computing Conference (IWCMC), (Istanbul, Turkey), pp. 820–825, 4-8 July 2011.
- [12] J. N. A.-K. Khalil M. Youcef and A. M. Shatnawi., "Intelligent traffic light flow control system using wireless sensors networks," Journal of information science and engineering, vol. 26, pp. 753–768, 2010.
- [13] G. Yang, "An improved vehicle-logo localization algorithm based on texture analysis," in International Conference on Computer Science and Information Processing (CSIP), (Xian, Shaanxi, China), pp. 648–651, 24-26 August 2012.
- [14] L. L. Li, W., "A novel approach for vehicle-logo locationbased on edge detection and morphological filter.," in Second International Symposium on Electronic Commerce and Security, (Nanchang, China), pp. 343–345, 22-24 May 2009.
- [15] G. Loy and J.-O. Eklundh, "Detecting symmetry and symmetric constellations of features.," in In Proceedings of the 9th European conference on Computer Vision, vol. Part II, (Berlin), pp. 508–521., Heidelberg: Springer-Verlag, 2006.
- [16] S. F. Stefan Badura, "Advanced scale-space, invariant,low detailed feature recognition from imagescar brand recognition," in Proceedings of the International Multiconference on Computer Science and Information Technology, (Wisa, Poland), pp. 19–23.,October 1820 2010.
- [17] C. N. E. A. Apostolos P. Psyllos and E. Kayafas, "Vehicle logo recognition using a sift-based enhanced matching scheme," IEEE Transactions on Intelligent Transportation Systems, vol. 11, pp. 322–328, June 2010.
- [18] D. G. Lowe, "Object recognition from local scaleinvariant features," in Proceedings of the International Conference on Computer Vision., vol. 2, pp. 1150–1157, 1999.
- [19] T. T. H. Bay, A. Ess and L. V. Gool, "Surf: Speeded up robust features," in computer Vision and image Understanding (CVIU), vol. 110, pp. 346–359, june 2008.
- [20] Z. Nan, "Vehicle-logo recognition based on classifier combinations," in 2012 3rd International Conference on System Science, Engineering Design and Manufacturing Informatization (ICSEM),, vol. 2, (Chengdu,China), pp. 278–281, October 2012.
- [21] D. G. Lowe, "Distinctive image features from scaleinvariantkeypoints," International Journal of Computer Vision, vol. 60, pp. 91–110, nov 2004.
- [22]Basma M. Mohammad El-Basioni, Sherine M. Abd El-kader, and Hussein S. Eissa, "Improving LLEAP Performance by Determining Sufficient Number of Uniformly Distributed Sensor Nodes", CiiT International Journal of Wireless Communication, Volume 4, Issue 13 , 2012, pages: 777–789.

- [23] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *Journal of the Optical Society of America*, vol. 7, pp. 923–932, 1990.
- [24] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, pp. 1–27, April 2001.
- [25] M.S. Sarfraz, A. Saeed, M. H. Khan, and Z. Riaz, "Bayesian prior models for vehicle make and model recognition," in *Proceedings of the 7th International Conference on Frontiers of Information Technology, FIT '09*,

Maha El Meseery gained Msc and Bsc. in computer engineering from Cairo University in 2009 and 2003 respectively. Since 2007, she works as a research assistant in Electronics Research Institute. Her research interests are image processing and pattern recognition.

Mahmoud Fakhr El-Dein was born in Giza, Egypt, in 1968. He received the B.S. degree in automatic control from the University of Minufia in 1991 and the M.Sc. and PhD degree in Computer Engineering, from Electronics and Communications Department, Cairo University, Faculty of Engineering, in 1996 and 2004 respectively. He works as an associate professor of computer science at the Electronics Research Institute. He has also worked at the Faculty of Engineering, Abha, KSA and different universities in Egypt. He works as a consultant at the ministry of communications and information technology, Egypt, 2004 – till now and collects different experiences in information technology while working in different governmental organizations in Egypt. His areas of interest are Evolutionary Computation, Advanced automatic control, image processing and biomedical engineering

Sherine. Abd El-kader has her MSc, & PhD degrees from the Electronics & Communications Dept. & Computers Dept., Faculty of Engineering, Cairo University, at 1998, & 2003. Dr. Abd El-kader is an Associate Prof., Computers & Systems Dept., at the Electronics Research Institute (ERI). She is currently supervising 3 PhD students, and 10 MSc students. Dr. Abd El-kader has published more than 25 papers, 4 book chapters in computer networking area. She is working in many computer networking hot topics such as; Wi-MAX, Wi-Fi, IP Mobility, QoS, Wireless sensors Networks, Ad-Hoc Networking, realtime traffics, Bluetooth, and IPv6. She is an Associate Prof., at Faculty of Engineering, Akhbar El Yom Academy from 2007 till 2009. Also she is a technical reviewer for many international Journals. She is heading the Internet and Networking unit at ERI from 2003 till now. She is also heading the Information and Decision making support Center at ERI from 2009 till now. She is supervising many automation and web projects for ERI. She is supervising many Graduate Projects from 2006 till now. She is also a technical member at both the ERI projects committee and at the telecommunication networks committee, Egyptian Organization for Standardization & Quality since February 2007 till 2011.

Online Newspaper using Mashup

Ramiz Kastrati¹, Mentor Hamiti²

¹ College Universum
Ferizaj, 70000, Kosovo

² South East European University
Tetova, 1200, Macedonia

Abstract

Online Newspaper has found a great use in all domains of life, whereas in Albanian speaking areas the development of online newspaper for the gathered information does not have a wide dimension. For this reason we have presented in our paper the methodology of gathering the information and its filtration from the scattered locations into a single one. By using Mashup technology, through Yahoo! Pipes application we will rank and filter the relevant information, which in this case are job offers. Filtered and selected information will be set in a blogger for distribution.

Keywords: Mashup, Yahoo pipes, blog, newspaper, online.

1. Introduction

The main aim of this paper is gathering the scattered information and their orientation into one location, so we can access them from a website and find them in filtered form and sorted in categories.

The users of this blog will be able to get the scattered information searched on internet faster and in a comfortable way and include the information directly by using scattered systems.

The paper is based on recent developments of the newspapers and electronic dynamic magazines.

2. Web 2.0

Web 2.0, is a phrase firstly created in 2004 from O'Reilly Media, an American media company [1]. It's the second web based generation, like social media sites, Wiki like websites, various communication instruments, and the systems similar with so called folksonomy, which enables the online cooperation.

Web 2.0 gives opportunity to Websites users to do more rather than just browse and find the information. They can

expand and improve the interactive facilities of the first Web 1.0 generation for building computer variant network platform, which allows the users to administer the software applications simply with one browser. The users may master data in Web 2.0 and exercise controlling those data. These sites can have "participatory architecture" that encourages users to valorize the application, while they are using it. This also offers great advantage compared to traditional websites which keep their users as followers. On the other hand, the content of those websites can be modified only from the administrator.

Web 2.0 websites are frequently characterized with a rich and user friendly interface, based on Ajax, Flex applications. The sites can also have social networks feature.

Web 2.0 converts the static platform into truly interactive platform. Instead of being simple downloader for clients, the users are able to uploading and create. Therefore the media is indeed convergent instead of fragmented.

Most tools of Web 2.0 websites are free programs which can replace entirely the traditional applications, for which users usually pay.

There are many Web 2.0 technologies, but the four most used are: blog, podcast, social network and wiki [2].

3. Blog

Often blogs are described and defined mostly as personal online diaries. This is true and is only one aspect of blogs. Web-bloggers are a type of a website. Blog is the abbreviation of the Web- log (network diary) [8]. Web-bloggers present an easy and comprehensive way for publishing all sorts of contents, such as: news, education, analysis, humor, personal observations and opinions.

Publishers can be groups, organizations and individuals. Web-blogs are destined for all types of people and for various purposes. Web-blogs can be very strict and formal, or extremely informal. They can exist in the internet or inside an intranet. Blogs contain personal thoughts, arguments, news or other multimedia documents such as photos, video and audio recordings [4].

Someone can raise the question on what is a blog, or how can we know that a particular site is a blog? The easiest way to find out the answer is to ask the blog creator. If the creator calls it a blog, then we also should accept it as such [4]. The reason for this is that there isn't a clear distinction between a blog and other forms of publishing. But, there is much coverage between each other.

Most creators do not consider online publishing as blogs even though most sites have blog characteristics.

4. Mashup

Mashup is a combination of two, or more data, presented and set in one tool [9]. Mashup can include text, graphics, maps, video, or other various forms of digital media. Usually internet applications also can be viewed through a web browser. Third parties often take the existing data and combine them, in order to include these data or create an all new meaning which is not in disposal by the individual data.

The term mashup was introduced into the everyday dictionary referring to the new created tunes by combining two or more existing ones. One data mashup has reached to describe the combined practice of two or more groups of electronic data for increasing the actual meaning, or creating a new meaning, where as a result the original data and the product are available on internet.

Mashup applications are issued by some of the major companies in the information industry, such as Google, Yahoo, IBM and Microsoft. Lately there is a considerable effort for creating tools that can allow users with no particular programming skills to create their own mashup [10]. An important feature of this kind allows the possibility for those internet users that don't have experience in programming, to use thousands of mashup that are in disposal.

5. Yahoo Pipes

Yahoo Pipes is a well-known, widely used visual programming application for creating data mashup by aggregating, manipulating, and publishing web feeds [11].

Large amount of information sometimes can be a double edged sword; especially when the amount of information increases so it becomes difficult to process and to follow it. Pipes use a flowchart approach for building mashup [12]. RSS atom feed and similar forms for information share have fastened the sharing to users, but dealing with information of that kind becomes an increasing problem. Yahoo has started lately a public solution of this problem through Yahoo Pipes service (pipes.yahoo.com).

A general view of yahoo pipes application is presented in the following figure:

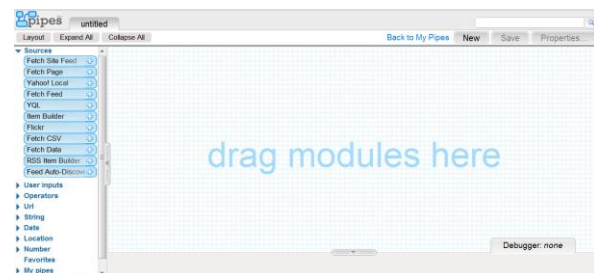


Fig. 1. Yahoo Pipes view and its modules.

Yahoo Pipes (<http://pipes.yahoo.com/>) is an interactive web application for data management. A yahoo pipe summarizes information from web sources by using a user friendly graphic interface. This is used for creating a Mashup web, by building a diagram where the information flows through various activity elements, used for analyzing and taking the process of information. This tool allows all users to search, as well as filtrate the actual data from internet sources, at any time.

Internet users are empowered with new abilities by shaping the content of the web, which presents a different thing in the web history. Users can share the created information and comments for norms, products, services and various notes made by companies and institutions, but also and those created by other users. Traditional phrase Web 2.0 symbolizes this change in the internet. In a passive way, users are inactive clients of information in disposal, managed by redactors and publishers of information, where the users create data for other users thus putting aside mediators. This tendency brought a need for new technical and organizational user oriented solutions that support this active role. The final result is the increase of the amount of available information online, and the problem of finding relevant information.

Term "pipes" comes from UNIX, meaning the way how applications communicate among themselves in data exchange [4]. In Yahoo Pipes Service, term "pipe" means the way information is combined from numerous sources.

A simple example is the possibility to combine information from numerous RSS feeds and to present them as a provider. Surely, Yahoo Pipes service offers more than this and allows users to combine various contents of Internet services for their needs.

Figure 2 presents the way data are combined from numerous RSS through Yahoo Pipes, and the information filtering.

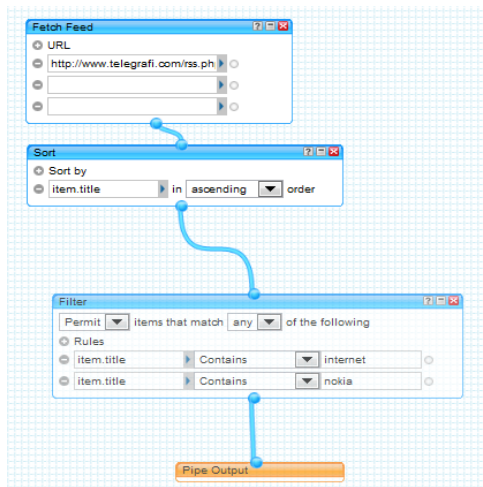


Fig 2. Usage of RSS and information filtration.

Fortunately, Web has done many innovations, so all users can fulfill their desires. Free services from Google and other companies allow organizing your job, sharing and synchronizing calendars, or creating maps for specific aims. You do not need to know what is API, and how to do JavaScript or XML programming, even though you can learn if you want.

It is understandable that Google is not the only “player” on the web. Here are some of our preferable services that allow us to create our sources for providing information, small and different programs as well, for applications where demands are limited only in your imagination.

A service that is offered by services for finding information is the one for finding locations by using Google Maps.

RSS is a good way for other users that read your content in the web. As RSS is supported almost in every browser, e-mail clients, web portals and search engines, and your site can be available for interested users on the subjects you write. RSS Feed is a file text that contains the title of your page and individual articles with URL addresses. When it comes to the simple sites, users can create photos manually, by using a text editor and RSS 2.0

specifications. On internet sites, RSS board can enter in values and variables for proving their source [5].

However, it is easier to use one of numerous automatic RSS generators, which can find HTML tags and generate an XML file. From tens of such services, it's preferred for starters to use FeedYes (feedyes.com), because do not only searches automatically for web sites, but it helps you use sources manually. When you do this, you can use feedvalidator.org to control your mistakes in RSS, or you can use a tool for controlling the place of RSS board. If everything is correct, the source will pass through a free service called FeedBurner (www.feedburner.com). When you are in FeedBurner site, you consider the source of information as separated by others through Google AdSense program.

News sources enable you to be updated with events, without taking your time. But, if you want to find detailed information, Yahoo Company enables it through Yahoo pipes service.

Yahoo Pipes is an online free service, that allows people without experience in programming to create applications and to use them (e.g. Really Simple Syndication [RSS]) [4] and to create Mashup of data by using a visual drag-and-drop editor. Yahoo Pipes also offers a library of pipes (momentarily enumerates with thousands), that enables users to copy, reuse and modify pipes for their needs. Yahoo Pipes can offer a more focused group that provides news, rather than a traditional combined provider offered from sites like Yahoo. Our web site in Yahoo allows users to choose news from imposed pre-projections, whereas Yahoo Pipes allows users to design data and information sources that are interesting for users in a particular moment, despite from the previous design of the owner.

Pipes also allow data supply analyses that previously were unreachable because of conditions such as: large amount of data or foreign language limitations [5].

Another unique attribute of pipes is the ability of including RSS feeds in the site of the Web publisher, without the need for prior specific software configuration on server-side.

6. Online Newspaper

For this paper we exploited applications from Yahoo Pipes, <http://pipes.yahoo.com> based on demands for creating an online newspaper, which will provide information for employment in Kosovo, Macedonia and Albania. Beside this we are concentrated in providing daily and technology information.

Figure 3 presents one of the employment applications searching engine which has generated a code as written below.

```
<script
src="http://l.yimg.com/a/i/us/ppp/listbadge_1.6.js">{"pipe
_id":"2e69414d2ad868c6b197494de4e75ea2", "_btype":"li
st", "pipe_params":{"Lokacioni":"","Puna":"programmer"}
}</script>
```

Below is the image of modules which we have used for employment application, set for various rankings. In our case, we have searched for a Programmer, Oracle, Teacher, and PHP by using mentioned modules in Yahoo Pipes.

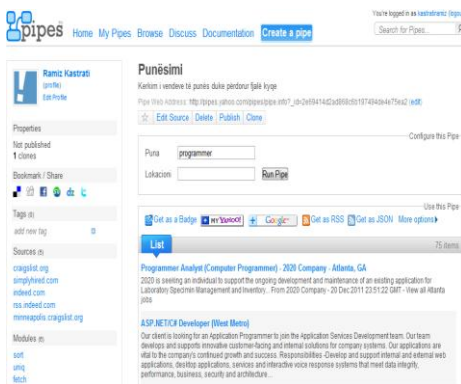


Fig 3. Image of an application for search engine.

The code below is generated for searching in the employment application in Kosovo, Albania and Macedonia.

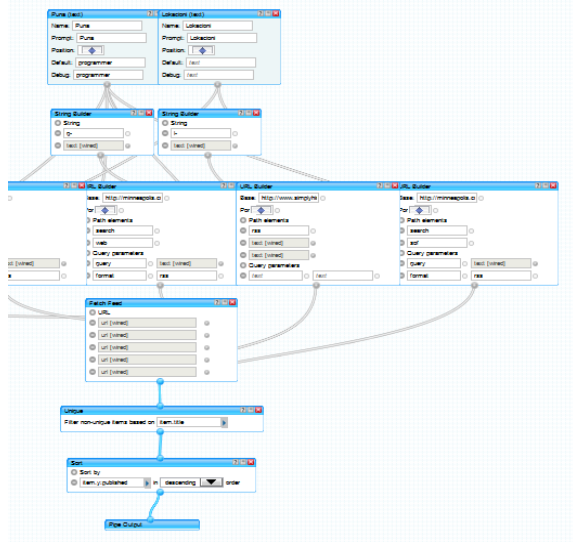


Fig 4. Modules created in Yahoo Pipes for employment by using search engine for jobs and location.

The code below is generated for searching in the employment application in Kosovo, Albania and Macedonia.

Whereas the application module for this search which uses various web portals is presented in Figure 5.

```
<script
src="http://l.yimg.com/a/i/us/ppp/listbadge_1.6.js">{"pipe
_id":"177d2d1b7e5661852cd5287a2a6ae7a", "_btype":"li
st"}</script>
```

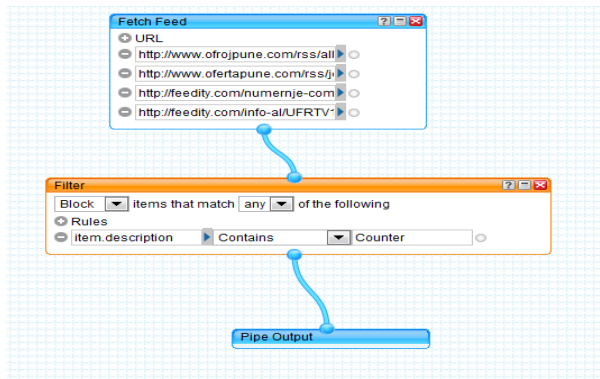


Fig 5. Application for selection from various employment web portals.

The generated code for information on Technology created from the application in Figure 6 is given below.

```
<script
src="http://l.yimg.com/a/i/us/ppp/listbadge_1.6.js">{"pipe
_id":"28a752db487c62bd0d319d090fc7105e", "_btype":"li
st"}</script>
```

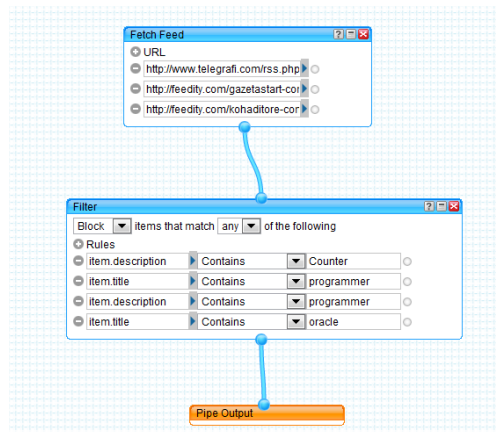


Fig 6. Application for filtering and finding information for Technology.

We also used RSS direct services to demonstrate their functionality for e.g.: For finding jobs which are presented

as gathered information according to given filters in Yahoo Pipes application.

Usage of Mashup technology and its application on blog is seen as an efficient and fast method for extracting information from various web sites and setting them in a particular place, in order to use them.

The created blog gives enough information thanks to this technology and its usage is satisfactory.

The image of the blog is given in Figure 7, whereas registration of this blog is done on www.blogspot.com which is a very appropriate platform that supports Mashup technology.



Fig 7. Finalized newspaper blog.

7. CONCLUSIONS

In this paper we have managed to gather in one particular place the needed information through Mashup technology with visual application of Yahoo Pipes. This information is set in blogger which provides users to filter the information according to their requirements. Various module usages which Yahoo Pipes offers, with gathered information offered for employment, or published from numerous websites, are usable in one particular place. Also we made categorization of employment offers. Furthermore, for completing a newspaper, we filtered daily news from portals in Albanian language in order to offer completed services.

From this paper we can conclude that this technology gives satisfactory results for those individuals that don't have enough knowledge to work with pure code, but they are willing to use information technology.

References

- [1] Matthew D. Dinmore, C. Curtis Boylls Empirically-Observed End-User Programming Behaviors in Yahoo Pipes
- [2] Jeff K. Wilson, Made-to-order content with Yahoo Pipes
- [3] Jody Condit Fagan, Mashups for the Nontechies: Yahoo Pipes
- [4] Yahoo Pipes, Available: <http://pipes.yahoo.com>, Accessed: 20.04.2013
- [5] Blooger, Available: <http://www.blogger.com/home>, Accessed: 10.04.2013
- [6] Blooger, Available: <http://blog.larkin.net.au>, Accessed: 24.03.2013
- [7] Pinkie Anggia and Dana Indra Senseuse, Analyzing the Benefits of Blog for Enhancing Knowledge, Management among Bloggers, IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 2, No 1, March 2013
- [8] Y. Wang, "Blog for Teacher Implicit Knowledge Sharing Model Construction", in International Conference of Information Technology, Computer Engineering and Management Sciences, 2011, Vol. 3, pp. 363-367.
- [9] Mashups with Yahoo Pipes, Available: <http://www.pcmag.com/article2/0,2817,2193042,00.asp> Accessed: 24.03.2013.
- [10] S. Peenikal, Mashups and the Enterprise: Whitepaper, Mphasis - an HP company, September 2009.
- [11] Matthew D. Dinmore, C. Curtis Boylls, Empirically-Observed End-User Programming Behaviors in Yahoo! Pipes, Asian Journal of Information Technology, 2011
- [12] Wang, G., Yang, S., and Han, Y., Mashroom: end user mashup programming using nested tables, 2009.

Ramiz Kastrati, holds a MSc diploma in Faculty of Contemporary Sciences and Technologies at South East European University (2012). Currently he is Assistant at the University College Universum, R. Kosovo.

Mentor Hamiti, holds a PhD diploma in Faculty of Contemporary Sciences and Technologies at South East European University (2010). Currently he is Assistant Professor at the same University, R. Macedonia (www.seeu.edu.mk).

Lock-Free Readers/Writers

Anupriya Chakraborty¹, Sourav Saha², Ryan Saptarshi Ray³ and Utpal Kumar Ray⁴

¹ Department of Information Technology, Jadavpur University Salt Lake Campus
Kolkata, West Bengal 700098, India

² Department of Information Technology, Jadavpur University Salt Lake Campus
Kolkata, West Bengal 700098, India

³ Department of Information Technology, Jadavpur University Salt Lake Campus
Kolkata, West Bengal 700098, India

⁴ Department of Information Technology, Jadavpur University Salt Lake Campus
Kolkata, West Bengal 700098, India

Abstract

The past few years have marked the start of a historic transition from sequential to parallel computation. The necessity to write parallel programs is increasing as systems are getting more complex while processor speed increases are slowing down. Current parallel programming uses low-level programming constructs like threads and explicit synchronization using locks to coordinate thread execution. Parallel programs written with these constructs are difficult to design, program and debug. Also locks have many drawbacks which make them a suboptimal solution.

Software Transactional Memory (STM) is a promising new approach to programming shared-memory parallel processors. It is a concurrency control mechanism that is widely considered to be easier to use by programmers than locking. It allows portions of a program to execute in isolation, without regard to other, concurrently executing tasks. A programmer can reason about the correctness of code within a transaction and need not worry about complex interactions with other, concurrently executing parts of the program.

This paper shows the concept of writing code using Software Transactional Memory (STM) and the performance comparison of codes using locks with those using STM.

Keywords: *Parallel Programming; Multiprocessing; Locks; Transactions; Software Transactional Memory*

1. Introduction

Generally one has the idea that a program will run faster if one buys a next-generation processor. But currently that is not the case. While the next-generation chip will have more CPUs, each individual CPU will be no faster than

the previous year's model. If one wants programs to run faster, one must learn to write parallel programs as currently multi-core processors are becoming more and more popular. The past few years have marked the start of a historic transition from sequential to parallel computation. The necessity to write parallel programs is increasing as systems are getting more complex while processor speed increases are slowing down. Parallel Programming means using multiple computing resources like processors for programming so that the time required to perform computations is reduced [1].

2. The Readers/Writers Problem

The Readers/Writers Problem of Synchronization can be described as follows:

An object is shared among many threads, each belonging to one of two classes:

- Readers: read data, never modify it
- Writers: read data and modify it

Using a single lock on the data object is overly restrictive:

- There are many readers reading the object at once
- Allow only one writer at any point
- We must control access to the object to permit this protocol.

Correctness criteria:

- Each read or write of the shared data must happen within a critical section.
- Guarantee mutual exclusion for writers.
- Allow multiple readers to execute in the critical section at once.

3. The Readers/Writers Problem using Locks

3.1 Description

The hardest problem that should be overcome when writing parallel programs is that of synchronization. Multiple threads may need to access the same locations in memory and if careful measures are not taken the result can be disastrous. If two threads try to modify the same variable at the same time, the data can become corrupt. Currently locks are used to solve this problem. Locks ensure that a critical section, which is a block of code that contains variables that may be accessed by multiple threads, can only be accessed by one thread at a time. When a thread tries to enter a critical section, it must first acquire that section's lock. If another thread is already holding the lock, the former thread must wait until the lock-holding thread releases the lock, which it does when it leaves the critical section [2].

In the Readers/Writers problem there are multiple readers and writers accessing the elements in the same buffer at the same time. The buffer is of fixed size. In the examples below we have taken the buffer size as 100000000. The problem is to synchronize these accesses properly so that when a write operation is occurring it should not be affected by any other read or write operation.

3.2 Reader and Writer Threads Code using Locks

The following code shows the reader and writer threads using threads and locks which solves the Readers/Writers problem:

```
void *reader(void *num_ptr1)
{
    unsigned char num1,*number_ptr1,
```

```
byte_under_stm1;
unsigned long k;
int i;
structtimevalini_tv;
number_ptr1=num_ptr1;
num1=*number_ptr1;
```

```
for((k=(((num1*ARRAY_SIZE)/(NUM_THREADS))));
k<(((num1+1)*ARRAY_SIZE)/(NUM_THREADS))/2;
k++)
{
    pthread_mutex_lock(&mutex);
    printf("The data read is %d\n",arr[rcount[num1]]);
    pthread_mutex_unlock(&mutex);
}
pthread_exit(0);
}
```

```
void *writer(void *num_ptr)
{
    unsigned long byte_under_stm1,k,ki=0;
    unsigned char num, *number_ptr;
    structtimevalini_tv;
    number_ptr=num_ptr;
    num=*number_ptr;

    for(k=(((num)*ARRAY_SIZE)/(NUM_THREADS));
k<(((num+1)*ARRAY_SIZE)/(NUM_THREADS))/2;
k++,ki++)
    {
        pthread_mutex_lock(&mutex);
        arr[rcount[num]]=1;
        printf("The data written is %d\n",
arr[rcount[num]]);
        rcount[num]++;
        pthread_mutex_unlock(&mutex);
    }
    pthread_exit(0);
}
```

3.3 Code Explanation

The above program was tested using NUM_THREADS ranging from 1 to 6 threads (in effect 12 threads are created, one for reading and one for writing, making up a total of 6 reader/writer pairs), created to access the values in array arr. There are two processes, reader and writer whose functions are respectively, as the names suggest. The array is divided into several parts depending on the

value of NUM_THREADS and the reader/writer pair accesses the corresponding part of the array.

The reader thread is invoked using the thread ID which is passed to it as the parameter *num_ptr*. Based on this parameter, each thread accesses the corresponding part of the array. Each element is first locked, thereby entering the critical section, where no other thread may access the data being read. Once the data is read, the lock is released. The writer process works in a similar manner. In the above program “reader” and “writer” are the two thread processes for reading and writing elements from the buffer respectively. Here the array *arr* is the buffer. The global array *rcount* keeps track of the position of elements in the buffer.

In the thread “reader”, elements are read from the buffer by the following statements.

```
for(k=(((num1*ARRAY_SIZE)/(NUM_THREADS)));  
k<(((num1+1)*ARRAY_SIZE)/(NUM_THREADS))/2;  
k++)  
{  
pthread_mutex_lock(&mutex);  
printf("The data read is %d\n",  
arr[rcount[num1]]);  
  
pthread_mutex_unlock(&mutex);  
}
```

In the thread “writer” elements are written into the buffer by the following statements.

```
for(k=(((num)*ARRAY_SIZE)/(NUM_THREADS));  
k<((num+1)*ARRAY_SIZE)/(NUM_THREADS)/2;  
k++,ki++)  
{  
pthread_mutex_lock(&mutex);  
arr[rcount[num]]=1;  
printf("The data written is %d\n",  
arr[rcount[num]]);  
rcount[num]++;  
pthread_mutex_unlock(&mutex);  
}
```

The following statement is used to record the time before the threads are created:

gettimeofday(&ini_tv,NULL);

The same call is also used to record the time when all threads have just finished their executions.

The total time taken is then calculated and printed using the following statement:

printf("Total Time Taken = %ld\n", final_tv.tv_sec - ini_tv.tv_sec);

As it can be seen from the above code snippet, 3 calls related with the mutex are being used. They are as follows:

- **pthread_mutex_init(&mutex,NULL)** is used for lock initialization.
- **pthread_mutex_lock(&mutex)** means that any thread must acquire the lock on mutex to execute the critical section following this function.
- **pthread_mutex_unlock(&mutex)** is used for unlocking.

In this program, the regions where more than one thread may access the global array, *rcount[]* at the same time are the critical sections. Thus these regions are enclosed within locks. Hence when a write operation is occurring in this program it is not being affected by any other read or write operation.

3.4 Experimental Results

We have taken all the experimental data for the outputs shown in this paper by running the codes on a machine which has 6 cores with hyper-threading. Thus, a maximum of 12 threads can run in parallel.

The experimental results for The Readers/Writers Problem using locks are presented below:

Number of Reader/Writer Pairs	Time Taken (Locks)
1	99
2	52
3	33
4	25
5	19
6	16

Table 1: Experimental Results for the Readers/Writers Problem using Threads and Locks

The above experimental data has been represented graphically in Figure 1 and Figure 2 which show the variation of Time Taken for execution of the code, and Speedup respectively, with increase in the number of

threads for the code of the Readers/Writers Problem using threads with Locks.

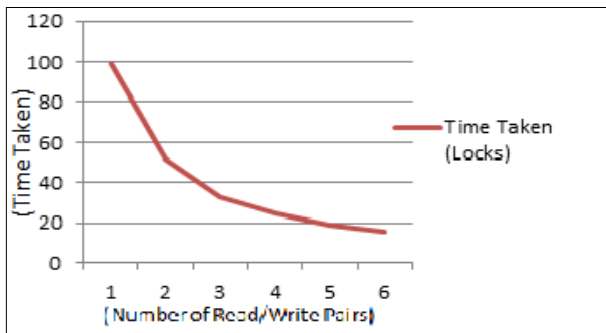


Fig. 1 : Graph showing the Time Taken vs. Number of Reader/Writer Pairs for the Readers/Writers Problem using Threads with Locks

It can be seen from the above graph that as the number of reader/writer pairs increases, the time taken for execution decreases.

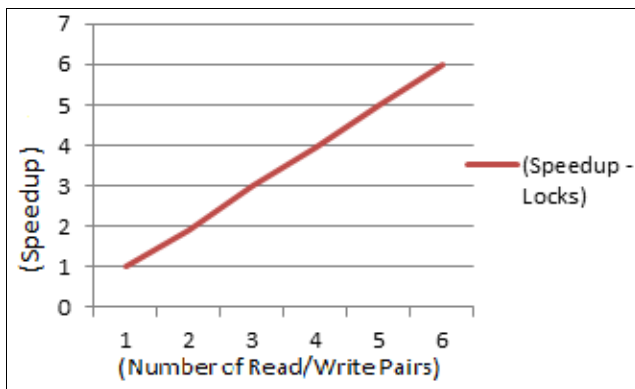


Fig. 2 : Graph showing the Speedup vs. Number of Reader/Writer Pairs for the Readers/Writers Problem using Threads with Locks

We can see that the speedup increases linearly with the number of reader/writer pairs.

4. The Readers/Writers Problem using STM

4.1 Description

The synchronization problem can also be solved using STM. If STM is used in a program then we do not have to

use locks in the program. Thus the problems which occur due to the presence of locks in a program do not occur in this type of code. The critical section of the program has to be enclosed within a transaction. Then STM by its internal constructs ensures synchronization in the program.

There are 8 categories of calls associated with STM which have been used in this program. They are as follows:

- **stm_init** is used to initialize the TinySTM library at the outset. It is called from the main thread before accessing any other functions of the TinySTM library.
- **stm_init_thread** is used to initialize each thread that will perform transactions. It is called once from each thread that performs transactional operations before the thread calls any other functions of the TinySTM library. In this program it is called from the threads **reader** and **writer**.
- **stm_exit** is the corresponding shutdown function for **stm_init**. It cleans up the TinySTM library. It is called once from the main thread after all transactional threads have completed execution.
- **stm_exit_thread** is the corresponding shutdown function for **stm_init_thread**. It cleans up the transactional thread. It is called once from each thread that performs transactional operations upon exit. In this program it cleans up the threads **reader** and **writer**.
- **START(0,RW)** is used to start a transaction. In this program it is used in the threads **reader** and **writer**.
- **COMMIT** is used to close the transaction. In this program it is used in the threads **reader** and **writer**.
- **byte_under_stm1=(unsigned char)LOAD(&rcount)** stores the value of **rcount** in **byte_under_stm1**. In this program it is used in the threads **reader** and **writer**.
- **STORE(&rcount, byte_under_stm1)** stores the value of **byte_under_stm1** in **rcount**. In this program it is used in the threads **reader** and **writer**.

4.2 Reader and Writer Threads Code using STM

The following code shows the reader and writer threads using threads and STM which solves the Readers/Writers Problem:

```
void *reader(void *num_ptr1)
{
    unsigned char num1,*number_ptr1;
    unsigned long k;
    unsigned long byte_under_stm1;
    struct timeval ini_tv;
    number_ptr1=num_ptr1;
    num1=*number_ptr1;

    stm_init_thread();

    for((k=(((num1*ARRAY_SIZE)/(NUM_THREADS))));
    k<(((num1+1)*ARRAY_SIZE)/(NUM_THREADS))/2;
    k++)
    {
        START(0,RW);
        byte_under_stm1=(unsigned char)
        LOAD(&rcount);
        printf("The data read is %d\n", arr[rcount[num1]]);
        STORE(&rcount,byte_under_stm1);
        COMMIT;
    }
    stm_exit_thread();
    pthread_exit(0);
}
```

```
void *writer(void *num_ptr)
{
    unsigned long byte_under_stm1,k,ki=0;
    unsigned char num, *number_ptr;
    number_ptr=num_ptr;
    num=*number_ptr;
    stm_init_thread();

    for(k=(((num)*ARRAY_SIZE)/(NUM_THREADS));
    k<((num+1)*ARRAY_SIZE)/(NUM_THREADS)/2;
    k++,ki++)
    {
        START(0,RW);
        byte_under_stm1=(unsigned char)
        LOAD(&rcount);

        arr[rcount[num]]=1;
    }
}
```

```
        printf("The data written is %d\n", arr[rcount[num]]);
        rcount[num]++;
        STORE(&rcount,byte_under_stm1);
        COMMIT;
    }
    stm_exit_thread();
    pthread_exit(0);
}
```

4.3 Code Explanation

The program structure of the above code is same as that of the program for readers-writers problem using threads and locks. The only difference is that STM is being used in this code.

In this program, the regions where more than one thread may access the global array *rcount[]* at the same time are the critical sections. Thus these regions are enclosed within transactions using STM. Hence when a write operation is occurring in this program it is not being affected by any other read or write operation.

4.4 Experimental Results

The experimental results for The Readers/Writers Problem using STM are presented below:

No. of Read/Write Pairs	Time Taken (STM)
1	116
2	59
3	36
4	25
5	23
6	18

Table 2 : Experimental Results for the Readers/Writers Problem using Threads and STM

The above data has been represented graphically in Figure 3 and Figure 4 which show the variation of Time Taken for execution of the code, and Speedup respectively, with increase in the number of threads for the code of the Readers/Writers Problem using threads with STM.

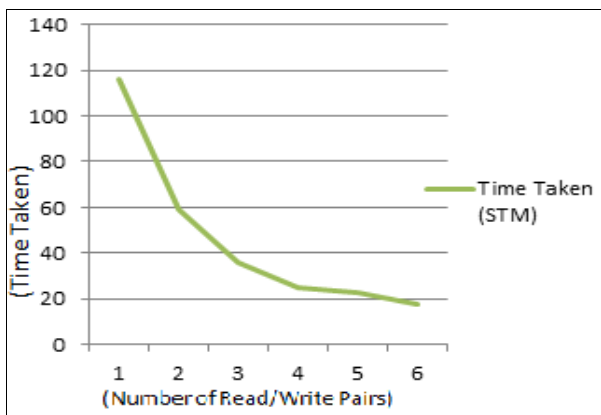


Fig. 3 : Graph showing the Time Taken vs. Number of Reader/Writer Pairs for the Readers/Writers Problem using Threads with STM

It can be seen from the above graph that as the number of reader/writer pairs increases, the time taken for execution decreases.

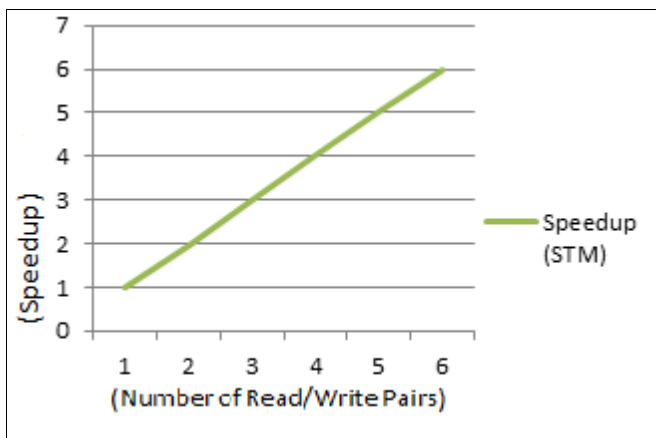


Fig. 4 : Graph showing the Speedup vs. Number of Reader/Writer Pairs for the Readers/Writers Problem using Threads with STM

We can see that the speedup increases linearly with the number of reader/writer pairs.

5. Conclusions

Figure 5 is a combination of Figures 1 and 3 as shown in previous sections.

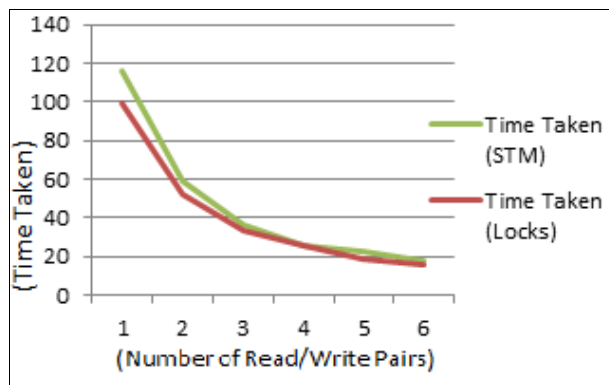


Fig. 5 : Graph showing the Time Taken vs. Number of Reader/Writer Pairs for the Readers/Writers Problem using both Threads with Locks and Threads with STM

We can see from the above graph that the time taken for executing the code using threads and STM is almost equal to the time taken for executing the same code using threads and locks. Thus, further research is being undertaken to improve the execution speed of STM.

Similarly, Figure 6 is a combination of Figures 2 and 4 as shown in previous sections.

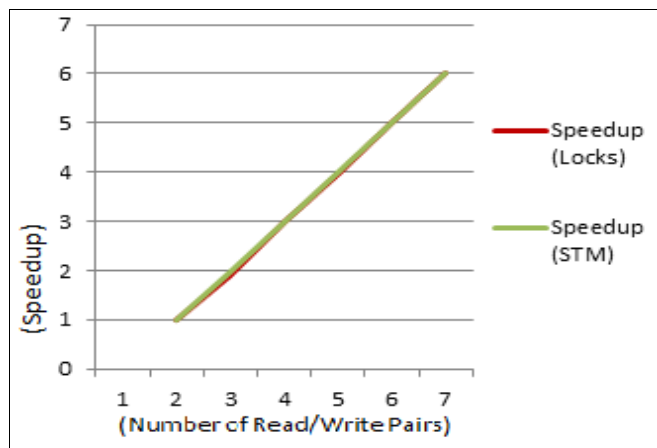


Fig. 6 : Graph showing the Speedup vs. Number of Reader/Writer Pairs for the Readers/Writers Problem using both Threads with Locks and Threads with STM

From the above graph we can see that the speedup for the codes using both locks and STM for this problem, are nearly the same. This is because the times taken for executing both the codes are also nearly equal.

STM has been shown in many ways to be a good alternative to using locks for writing parallel programs. STM provides a time-tested model for isolating concurrent computations from each other. This model raises the level of abstraction for reasoning about concurrent tasks and helps avoid many parallel programming errors.

This paper has discussed how STM can be used to solve the problem of synchronization in parallel programs, and in particular, the Readers/Writers Problem of Synchronization has been solved with Lock-Free Code using STM. STM has ensured that lock-free parallel programs can be written. This ensures that the problems which occur due to the presence of locks in a program do not occur in this type of code.

Many aspects of the semantics and implementation of STM are still the subject of active research. While it may still take some time to overcome the various drawbacks, the necessity for better parallel programming solutions will drive the eventual adoption of STM. Once the adoption of STM begins it will have the potential to pick up momentum and make a very large impact on software development in the long run. In the near future STM will become a central pillar of parallel programming.

References

- [1] [1] Simon Peyton Jones, "Beautiful concurrency".
- [2] Elan Dubrofsky, "A Survey Paper on Transactional Memory".
- [3] Pascal Felber, Christof Fetzer, Torvald Riegel, "Dynamic Performance Tuning of Word-Based Software Transactional Memory".
- [4] http://en.wikipedia.org/wiki/Transactional_memory
- [5] James Larus and Christos Kozyrakis. "Transactional Memory"
- [6] Pascal Felber, Christof Fetzer, Patrick Marlier, Torvald Riegel, "Time-Based Software Transactional Memory"
- [7] Tim Harris, James Larus, Ravi Rajwar, "Transactional Memory"
- [8] Mathias Payer, Thomas R. Gross, "Performance Evaluation of Adaptivity in Software Transactional Memory"
- [9] Kevin E. Moore, Jayaram Bobba, Michelle J. Moravan, Mark D. Hill, David A. Wood., "LogTM: Log-based Transactional Memory"
- [10] Dave Dice , Ori Shalev , Nir Shavit., "Transactional Locking II"
- [11] <http://tmware.org>
- [12] Maurice Herlihy, J. Eliot B. Moss, "Transactional Memory:Architectural Support for Lock-Free Data Structures".
- [13] Martin Schindewolf, Albert Cohen, Wolfgang Karl, Andrea Marongiu, Luca Benini, "Towards Transactional MemorySupport for GCC".
- [14] Virendra J. Marathe, Michael F. Spear, Christopher Heriot,Athul Acharya, David Eisenstat, William N. Scherer III, Michael L. Scott, "Lowering the Overhead ofNonblocking Software Transactional Memory".
- [15] Utku Aydonat, Tarek S. Abdelrahman,Edward S. Rogers Sr., "Serializability of Transactions inSoftware Transactional Memory".
- [16] Maurice Herlihy, Nir Shavit, "The Art of Multiprocessor Programming".
- [17] Brendan Linn, Chanseok Oh, "G22.2631 project report: software transactional memory".
- [18] Ryan Saptarshi Ray , "WritingLock-FreeCode using Software Transactional Memory".
- [19] http://en.wikipedia.org/wiki/Software_transactional_memory
- [20] <http://research.microsoft.com/~simonpj/papers/stm/>
- [21] http://www.haskell.org/haskellwiki/Software_transactional_memory.

First Author is currently a 3rd year student of B.E. (I.T.) in Jadavpur University, Kolkata, West Bengal, India.

Second Author is also a 3rd year student of B.E. (I.T.) in Jadavpur University, Kolkata, West Bengal, India.

Third Author is a research scholar in the Department of Information Technology, Jadavpur University, Kolkata, West Bengal, India.

Fourth Author is an Associate Professor in the Department of Information Technology, Jadavpur University, Kolkata, West Bengal, India.

The Utilization of Internet Communication into Knowledge Sharing Activities in Primary Schools

Lokman Mohd Tahir¹, Mohd Fadzli Ali², Tina Abdullah³, Mehmet Ozay⁴, Haswadi Hassan⁵, and Khadijah Daud⁶

^{1,4 & 6} Department of Educational Foundations, Universiti Teknologi Malaysia, Malaysia

² Department of Educational Technology, Universiti Teknologi Malaysia, Malaysia

³ Language Academy, Universiti Teknologi Malaysia, Malaysia

⁵ Department of Computer Science, Universiti Teknologi Malaysia, Malaysia

Abstract

The transformation of the Malaysian national policy from an industrial-based to a knowledge-based nation has prompted educational institutions to adopt a culture that promotes the advancement in knowledge, skills and area of expertise among the academic staff in schools. This study identifies the knowledge sharing activities using internet among teachers in the national schools. In addition, it also examines the relationship between environmental factors which consisted of technological factors, leadership, culture, incentives and school management system and the knowledge sharing culture. Apart from that, factors that impede the development of knowledge sharing in schools, which include teacher's attitude, time and workloads, were also studied. Quantitative analyses revealed that the teachers were practicing knowledge sharing with the wide usage of internet as part of their professional development activities and regression analysis also proved that school leadership was perceived to be the dominant factor that influences the development in knowledge sharing culture in educational institutions.

Keywords: internet communication; technology in education, knowledge sharing, knowledge management

1. Introduction

The importance of knowledge sharing and producing knowledge worker has taken a dramatic increase in organizational and human resource development issues recently [36] due to implications towards organizational effectiveness [2]. The emergence of knowledge as a tool towards achieving organizational vision started when the technological advancement in the information and communication technologies occurred with the advent of computer and the Internet. The Internet provides an effective communication infrastructure and creates new forms of human interactions to enable knowledge sharing through communication software and hardware. It is this innovative way of storing, accessing and distributing information that has transformed society and

organizations [1]. Soon the world is a boundary less space and opportunity and many organizations, be it private or otherwise, have thrived successfully. Therefore, it is with that in mind, Malaysia considered knowledge as an important advantage for the development of individuals and businesses as it encourages effectiveness in organizations and economic growth in the nation. Gradually, economic policies in Malaysia shifted its aims from an industrial-based to knowledge-based (K-Economy) [6] through the wide usage of computer-assisted systems. Soon after, government officials and organizational leaders were encouraged to inculcate the paradigm shift in their routine teaching tasks.

2. Literature review

2.1 Knowledge sharing: definition and issues

The concept of integrative knowledge-based schools which is defined as a process of understanding and appreciating knowledge with respect to differentiated functions and depending on the effective processes and knowledge sharing activities among members of any organization [24], [31]. [49]. On the other hand, it is argued that knowledge sharing deals with a process of telling, interpreting and educating other colleagues through the process of sharing information whether internally or externally through ICT-based assisted approaches in realizing organizational vision, mission and objectives by a requirement on how the teaching staff absorb, sharing and applying the knowledge and information on their working environment. Knowledge sharing concept is also defined as a set of behaviors and exchange of information or assistance to the other members in the same organization [11]. The significant impact of the knowledge sharing through Internet communication creates the concept of community of practice. This concept was introduced by Lave and

Wenger (1998) [32] postulate the idea of people with a common interest in the same area sharing ideas, developing innovations and collaborating over a period of time. In the context of this paper, knowledge sharing is mainly viewed from teachers' perspective in order to improve their teaching skills. This includes both formal and informal knowledge sharing. Collaboration through knowledge sharing also allows teachers to co-ordinate different types of teaching approaches and to avoid any duplication [22].

2.2. Knowledge Sharing Using Internet

Theoretically, ICT has significant impact towards improving teacher training and professional development especially to improve capabilities in teaching and learning processes [13]. In most EU countries, teachers are much more committed to the use of internet with flexible and adaptable system that caters the needs of teachers [23] aligned with EU policies of creating a knowledge-based economy [19]. As an educator in a digital and knowledge era, teachers are urged to work in networks, take part in networks, and consider that their role is to develop human networks or other colleagues for learning. Being part of a network, being able to behave in a network, taking benefits from networking demand teachers themselves experience such networked activities [17]. Beneficial aspects of using ICT in teachers training are remarkable based on teacher's responses towards the usage of internet in knowledge sharing activities; as a part of informal teacher training and development. An internet-based training is considered as flexible and adaptable approach to cater teachers' needs while minimizing travel costs and increasing internet competences [7]. In Germany for example, internet communication through the application of e-learning is considered as suitable learning tools to improve teachers and student teachers capabilities in teaching in creating positive networks that play as a tool for competence development and knowledge exchange among teachers because teachers are well-developed with self-management skills and motivated to use internet individually through usage of internet for sharing and development of knowledge and changing and sharing their experiences [26].

In Malaysian setting, application and digital technologies enhancement in education were always placed as top priority by the Malaysian government under the Ninth Malaysia Plan from 2006 to 2010 [33]; [43]. The Ministry of Education through its strategic policy on ICT revealed three vital objectives related to the internet communication implementation in schools. The first policy was that of ICT is used as an enabler to reduce the digital gap between the schools. The second, ICT plays a

role and function in education as a teaching and learning tool, as part of a subject and as a subject by itself. Thus, this policy emphasizes the importance of having access to information, communication, and as a productivity tool via computer technology. ICT as a part of a subject refers to the use of a software in each respective subject such as "invention" and "engineering". The third policy has a direct relation with using ICT so as to improve productivity of the management system. Based on the implementation effort, the Ministry, namely the Division of Educational Technology was assigned to materializing the following multi-phase strategies to make sure that the objectives of ICT will be achieved. Listing from the five important pillars, the educational technology division is committed to developing of ICT-based knowledge and skills among the students and teachers. By this implementation objective, ICT and internet communication will be extensively used to automate and mechanize work processes such as the processing of all related documentations such as, generating timetables, lesson planning, management of information systems and financial management [14]; [37].

The setting up of the ICT and internet connectivity in local school systematically will enhance the platform of the knowledge sharing system among teachers through the use of e-mails, blogs and school web pages as tools for sharing knowledge, information and disseminate practical skills to other teachers. With the objective of creating knowledgeable society and workforce, utilizing techno-education as their culture they culminate in utilization by others. In addition, they also realize the creation of knowledge-society to intensify the usage of information and communication technology in schools [14]. The andragogical learning and teaching process for enhancing teaching professional development through the development digital enhancement lead to the country's education reform and revamp in the near future to transform the Malaysian teachers into knowledge society and a developed country.

The role of ICT and internet communication in Malaysian schools is to connect and wire other teachers as a cost effective move in disseminations of knowledge and learning to other teachers with the objective to overcome the hurdles with maximizing the vast and richness under the rapid technological and globalization of information. Salbiah (2003) [43] pointed out that the use of internet in Malaysian local schools is to improve the learning process and, in particular, on interactive teaching and learning in which students and teachers have to collaborate with their partners and colleagues from other schools. Collegial interaction within and between schools using internet facilitates such as e-mail automatically will enhance the sharing of resources, expertise and advices and

information related to teaching and learning among teachers. This remarkable effort would enable them to be centered into one place not only for the purpose of networking among their peers but also for social connectivity and values such as friendship, racial integration, religious tolerance and knowledge. The rapid development of computers and information technologies in Malaysian educational system is relevance to support and play as a mediator for knowledge sharing, develop their teaching skills, gaining and sharing knowledge with other teachers as a part of the system that supports exchange and socialization of knowledge [3]. Using ICT that encompasses social networking, e-learning and blog spots as tools for knowledge sharing, makes teachers to be able to have an access, share, analyze, and disseminate information gained from a variety of sources such as blogs, web pages and e-mails in order to improve their teaching and knowledge related to professionalism issues. By the usage of internet, teachers have more chances to involve in an active working environment in collaborative and independent manner. On the basis of their experiences, teachers have some insights to choose appropriate educational software which creates more choices and allows transferences of skills and stimulate thinking skills. In fact, the function of internet communication within the scope of knowledge sharing is both to improve drastically the learning experiences of students and also to guide teachers to generate avenues for themselves as community practice member to work in a group and improve their team spirited attitudes, cohesion and social values. Implicitly, knowledge sharing system maintains the process through which both explicit and implicit knowledge is shared by other teachers. They take part in the system by supporting exchange, means, sharing of explicit knowledge, and socialization, say, promoting of tacit knowledge [39] among their colleagues.

3. The context of the study

Numerous literature have studied the knowledge sharing activities with the assistance of internet among teachers based on various setting that made this study a replicate version. However, least study was implemented using primary school teachers as their respondent that made this study significant to the contribution of teacher training research that related to the implementation of knowledge management in primary school setting. This research outlined four essential questions: (1) Do teachers implement knowledge sharing activities using ICT at school? (2) What factors influence the implementation of knowledge sharing activities using ICT and internet communication? (3) What are the main factors that

influence the implementation of knowledge sharing in schools using internet communication?

4. Methodology

4.1. Sampling

Research participants were selected from pool of primary school teachers that used ICT and internet as their medium for knowledge sharing activities whether formal and informally. Empirically, teachers that participated in this study as the research population came from the list of primary schools that were active in the ICT activities managed by the state department of education. The sampling data was provided by the State Department of Education with a total Of 300 teachers were listed as research sample for this study. A total of 300 teachers were selected through random sampling approach were contacted through their personal emails and later followed by an official letter to notify them that they being selected as sample for this study. Thus, 300 copies of the survey questionnaire were distributed to teachers who were contently willing to take part as respondents for this study. Out of 300 questionnaires that were distributed, only 200 sets of usable questionnaire were returned to the researchers, yielding a 76.5 percent response rate. The anonymity of teachers was protected with disclosing their real identities. The survey questionnaire was answered by participants based on consensus and voluntary basis and will be used as meaningful data in analyzing the validity and reliability of the measurement scales.

4.2 Instrumentation

The self-construct survey questionnaires have two major sections. Section one detailing the teachers' reflection related to the implementation of knowledge sharing in primary schools that comprised of four constructs which were the importance of knowledge sharing for teachers (6 items); positive social interactive through knowledge sharing using internet (6 items); teacher readiness in sharing knowledge with the assistance of internet in schools (6 items); improvement of teachers' potential, innovation, motivation and creativity improvement and job performance through Knowledge Sharing using Internet communication (6 items). The second section of the questionnaire was designed to measures the influencing (organizational leadership; technological and internet supports; institutional culture; recognition and rewards; management system) with 20 items and 5 items for each influencing factors. 12 items were constructed to measures the constraint (teachers' awareness and attitudes; time and workloads) factors with the

distribution of 4 items for each constraint factor. The overall total numbers of items studied were 56 items. The instrument was based on Likert Scale level of agreement consists of 56 items designed to obtain the teachers' perspective and evaluation on the influencing and constraint factors on the successful implementation of knowledge sharing using ICT in primary schools. All items were scaled using a five-item Likert scale ranging from 'strongly disagreed' (1) to 'strongly agreed' (5). In order to ascertain the content validity, panel of lecturers from the knowledge management, educational technology program and the state education officer was approached to assess content of the survey items. Based on the suggestions and constructive comments, some items were revised and changed. Next, instrument was piloted to 19 primary school teachers to determine the understandability of the items included in the questionnaire and also to incorporate any useful suggestions that they might offer. The item modification including rephrasing and rewording were done based on the feedback obtained from the pilot study. All items in the questionnaire were measured using the Cronbach alpha for overall internal consistency was at 0.8731 which considered as high reliability of measurement. Rasch analysis was used to determine the validity and reliability of the items studied. Reasons for resorting to Rasch analysis was empirically based into account the categorical and ordinal nature of such data [50] and small numbers of respondents that involved in the pilot study which were less than 20 respondents. Before proceeds with the actual data, rasch analysis was used to determine the reliabilities and the fit of the items. Item reliability that represents the dimensions of implementing the knowledge sharing through internet was measured at 0.98; while separation index of 0.94 showed the overall fit between the data and the model was good. Rasch analysis also indicated that constraint factors were determined at 0.92 and influencing factors were measured at 0.88 while separation index for items shows value ranging from 2.35 to 11. 56 items indicate acceptable values of more than 2.00. Hence, the data based on the pilot study was accepted based on the respondents' reliability index and item which indicates more than 0.8 is acceptable [50]. In addition, the separation index of more than 2.0 is also acceptable. The next step was determination of point measure correlation based on the summary of point measure correlation (PTMEA CORR) table for 56 items. All items show positive values with index > 0.30 which shows that items that were carefully constructed. Data was also evaluated based on the item and person measure is based solely on the logit scale. The results confirmed that all items are scattered and pointing towards the capability level of respondents' diversity. The item which is difficult

to be agreed upon is d7 with difficulty to be measured is 3.68 logit on the top scale, whilst the simplest item to be agreed upon is item a1 with measurement of -1.18 logit on the lower scale [51]. Rigorous evaluation showed that the achievement of the respondent or person is higher than the items in the instrument.

5. Results

This part describes the findings in research according to the research objectives as shown earlier. A total of 200 (66%) responses were Table 1 depicts the descriptive analysis. Generally, the teachers rated the implementation of knowledge sharing using internet in their schools as well implemented and was considered important to improve teachers' professionalism (mean score – 4.49). It is assumed that teachers place a greater priority and support on the implementation of knowledge sharing in their schools with the help of internet based instrumentation. Apparently, the teachers have been practicing knowledge sharing using internet internally and externally as part as their professional development activities with colleagues and other major subject teachers elsewhere. Empirically, teachers also believed that the implementation of knowledge sharing with ICT is considered as an important process and approach for teachers to enhance their performance, productivity and knowledge related to teaching specification and upgrading their major subject knowledge. In addition, teachers were perceived as active practitioners of the knowledge sharing processes which included the implementation of absorbing, sharing and applying the knowledge sharing approach while nurturing their career path.

Table 1: Means score of the implementation knowledge sharing using ICT.

<i>Constructs</i>	<i>Means</i>	<i>SD</i>
The importance of Knowledge Sharing using internet communication	4.85	0.39
Positive Social Interaction through Knowledge Sharing using internet	4.58	0.59
Readiness to Share Knowledge using internet	4.21	0.71
Improvement of Potential, Innovation, Motivation and Creativity Improvement and Job Performance through Knowledge Sharing using internet.	4.34	0.59
Total Mean Score	4.49	0.54

Additionally, this study also investigated the environmental factors which were considered influential in promoting the knowledge sharing culture in the identified educational institutions. Based on the literature

review, influencing factors such as technological and ICT supports, organizational leadership, institutional culture, recognition and incentives, and management system do influence development of the knowledge driven culture using ICT as knowledge catalyst. In contrast, factors that impede the implementation of knowledge sharing were teacher’s attitudes and awareness, time constraint, and academic workload. Table 2 below illustrates the list of influencing and constraining factors.

Table 2: Means score of influence and constraint factors

<i>Factors</i>	<i>Mean scores</i>	<i>SD</i>
<i>Influencing</i>		
Technological and ICT supports	4.16	0.68
Organizational leadership	4.37	0.54
Institutional culture	4.27	0.64
Recognition and incentives	3.97	0.73
Management system	3.95	0.62
<i>Constraining</i>		
Teacher’s attitudes and awareness	4.12	0.68
Time constraint	3.53	0.54
academic workload	4.27	0.65

Table 2 indicates factors that advocate and limit the knowledge sharing implementation using internet as mediator in primary schools. Based on the descriptive analysis, teachers selected organizational leadership as the most important determinant in promoting the knowledge sharing proven by the highest mean score of 4.37. The second highest noted was supportive culture (mean score - 4.27) followed by ICT and technology support (mean score - 4.16). Meanwhile, recognition and incentives (mean score - 3.97) and management system (mean score - 3.95) were considered as weak factors that support the implementation of knowledge sharing. In contrast, three important constraint factors were identified were teacher’s attitude and awareness, time constraint and academic workload. Based on the descriptive analyses, academic workload (mean score – 4.27) was chosen as the main constraints for teachers to share their knowledge with other colleagues followed by the teacher’s attitudes and awareness (mean score – 4.17) and lastly time constraint factor (mean score – 3.53).

Regression analysis listed the influencing factors, which were ICT and technology support, organizational leadership, institutional culture; recognition and incentives and the management system.

Table 3: Regression coefficients of teachers’ knowledge sharing implementation using ICT.

<i>Influencing Factors</i>	<i>t-values</i>	<i>Betas</i>	<i>p-values</i>
ICT and technological supports	1.63	0.13	0.11

Organizational leadership	3.91	0.27	0.00*
Institutional culture	2.71	0.24	0.01*
Recognition and incentives	-0.22	0.02	0.82
Management system	1.69	0.11	0.09

$p < 0.05$; ** $p < 0.01$

From the regression analysis, it is clear that knowledge sharing implementation using internet communication was strongly predicted by two important factor which were the organizational leadership [$t=3.91$, $\beta=0.27$, $p < 0.05$] based on teachers perception and evaluation towards the implementation on knowledge sharing using ICT in primary schools. In addition, institutional culture was also measured as the second important predictor for implementation on knowledge sharing [$t=2.71$; $\beta=0.24$, $p < 0.05$].

6. Discussion

It is apparent from this study that teachers’ perception provides some significant findings related to the concepts and implementation of knowledge sharing in an educational setting. As an educational organization that enhanced and promoted the concept of learning organization, the realization of knowledge sharing was particularly implemented through professional development approach in which teachers are urged to share their knowledge and skills internally with other teacher’s colleague especially novice teachers and externally with major subject teachers from other schools in order to produce highly capable, intellectual and skillful teachers.

According to the findings, it is presumed that teachers have high confidence towards the implementation of knowledge sharing blended with internet communication as the main catalyst and considered as an important asset for them to increase their teaching, classroom management and administrative skills in nurturing their capabilities in major teaching subjects and internet practical skills as part of informal teaching education and training that located in their schools. These findings reveal that teachers are practicing the concept of ‘community of practice’ (CoP) which is defined as a group of people who share, concern and deepen their knowledge through interactions among members in their organization [32] [29] and [30] with ICT as the main mediator for transferring and gaining the knowledge. The implementation of knowledge sharing truly promotes teachers to partake in discussions, contribute each other, and share their knowledge using e-mails, internet and intranet as instruments in delivering knowledge through

boundary less environment. This process is crucial in teacher professional development that improving their credibility in teaching. The above results generally confirm previous findings that showed significant knowledge sharing practices that used internet communication as their main catalyst in sharing teachers' knowledge and information without limited boundary and internet communication were considered as the important tools in realizing their networking in Malaysian context. In terms of measuring the relationship between influencing factors with knowledge sharing approach, it is indicated that there was a strong association between the influencing factors with knowledge sharing implementation using internet communication as catalyst in educational institutions. Based on regression analysis, the findings also showed that organizational leadership plays a salient role in supporting knowledge sharing environment among teachers. Organizational leader plays as an active catalyst or provider in promoting and enhancing sharing of knowledge among teachers especially using internet communication as the provider to achieve effective learning culture in educational organization. This is indeed a dimension worthy of another study. Nevertheless, this finding concurs with the study findings by Low (2000) [34] that indicate the most influential factor of knowledge sharing was organizational leadership. A study by Pan & Scarbrough (1999) [40] also found that the role of organizational leadership is a crucial factor in the implementation of knowledge sharing. This study also proves that there is a significant relationship between institutional cultures: staffs' attitudes and supports from group members, with the knowledge sharing implementation. In this sense, the improvement of the implementation of knowledge sharing in any organization also depends on staff attitudes and supporting climate from their colleagues

7. Conclusion

Finally, undoubtedly leadership is an important indicator in determining the successful of the knowledge sharing using the internet and ICT as main supporter whether in every institution that promotes learning as part of professional development process. Thus, organizational leaders should promote the culture of knowledge sharing using ICT as a medium to share knowledge and skills to improve human resource competencies. Thus, a leader determines the achievement of the implementation of knowledge sharing in organizations. He or she acts as a catalyst and control the activities in learning organizations. We see many examples of leaders who bring their organizations towards success and many

schools that implement the knowledge sharing ethos in their schools that reaps the rewards and accolades.

References

- [1] Abdullah and Hema Date (2009). Public sector knowledge management: a generic framework. Public Sector ICT Management Review. January to June. 3(1). Retrieved October 24 from [http://apps.intan.my/psimr/vol3no1%, Public%20Sector%20Knowledge%20Management](http://apps.intan.my/psimr/vol3no1%20Public%20Sector%20Knowledge%20Management)
- [2] Afiouni, f. (2007). human resources management and knowledge management: a road map toward improving organizational performance. *Journal of American Academy of Business, Cambridge*; Sep 2007; 11(2); Retrieved April 28, 2011 from Abi/inform Global Database, St. Thomas University Electronic Library Database
- [3] Aini Aripah and Norizan, M.Y. (2008) Using Teaching Courseware to Enhance Classroom Interaction as a method of Knowledge Sharing. *Journal of Information Systems, Research & Practices*, 1(1), 1 – 13.
- [4] Al-Alawi, A. I., Al-Marzooqi, N. Y. & Mohammed, Y.F. (2007). Organizational culture and knowledge sharing: critical success factors. *Journal of Knowledge Management*, 11(2), 22- 42.
- [6] Azmizam Abdul Rashid, Hamzah Jusoh & Jalaludin Abdul Malik (2008). The implementation of E-Government in local authorities towards enhancing urban government efficiency. Paper presented at Paper presented at International conference on social, development and environmental studies, The National University of Malaysia, 18-19 November
- [7] Bernad, D. (2010). Teacher Development Policies and Programs. ICT in Teacher Education: Policy, Open Educational Resources and Partnership. Proceedings of International Conference of IITE-2010. November 15–16, St. Petersburg, Russia.
- [8] Billig S. H., Sherry, L., & Havelock, B. (2005). Challenge 98: sustaining the work of a regional technology integration initiative. *British Journal of Educational Technology*, 36(6), 987-1003
- [9] Briscoe, C. & Peters, J. (1997). Teacher collaboration across and within schools: Supporting individual change in elementary science teaching, *Science Education*, 81(1), 51- 65.
- [10] Brockmeier, L. L., Sermon, J. M., & Hope, W. C. (2005). Principal's relationship with computer technology. *NASSP Bulletin*, 89(643), 45–63.
- [11] Wang, S., & Noe, R. A. (2010). Knowledge sharing: A review and directions for future

- research. *Human Resource Management Review*, 20(XX), 115–131.
- [12] Carroll, J., Choo, C., Dunlap, D., Isenhour, P. Kerr, S. Maclean, & A., Rosson, M. (2003) Knowledge Management Support for Teachers. *Educational Technology Research and Development*, 51(4), 42– 64
- [13] Cattaneo, A., & Boldrini, E., (2006). A Qualitative Research Pattern for the Surveying of a Competence Profile of the Teacher Using ICTs in a Blended Learning Project. Research papers on Competence Development in Online Distance Education and E-Learning. Selected Papers from the 4th EDEN Research Workshop in Castelldefels/Spain, October, page : 25-28.
- [14] Chan, Foong-Mae, (2002). Seminar/Workshop on the Promotion of ICT Education to Narrow the Digital Divide, 15 – 22 October 2002, Tokyo, Japan. Educational Technology Division, Ministry of Education
- [15] Chang, I.-H., Chin, J. M., & Hsu, C.-M. (2008). Teachers' Perceptions of the Dimensions and Implementation of Technology Leadership of Principals in Taiwanese Elementary Schools. *Educational Technology & Society*, 11 (4), 229 – 245.
- [16] Chong Siong Choy & Choi Yong Suk (2005). Critical Factors in the Successful Implementation of Knowledge Management, *Journal of Knowledge Management Practice*, 14 – 22.
- [17] Cornu, B. (2010). Digital Natives in a Knowledge Society: New Challenges for Education and for Teachers. ICT in Teacher Education: Policy, Open Educational Resources and Partnership Proceedings of International Conference IITE-2010. November 15–16, St. Petersburg, Russian Federation.
- [18] Darling-Hammond, L. & Ball, D. L. (1998). Teaching for high standards: What policymakers need to know and be able to do (Consortium for Policy Research in Education), National Commission on Teaching & America's Future
- [19] George Georgiadis, Menelaos Sotiriou, Evangelos Tsiopoulos, Alina Konstantinidi, Costas Tsolakidis, Pavlos Koulouris, Michalis Orfanakis, Sofoklis Sotiriou, Anna Touloumakou and Ellinogermaniki Agogi Elena Tavlaki (2009)., E-training teachers in rural areas through satellite. Selected Papers from the 5th EDEN Open Classroom Conference: Addressing eLearning, Innovation and Quality in Schooling. 5-7 October, Poitiers, France
- [20] Gupta, B., Iyer, L. S. & Aronson, J. (2000). Knowledge management: practices and challenges. *Industrial Management and Data Systems*, 100(1), 17-21.
- [21] Granger, C., Morbey, M., Lotherington, H., Owston, R., & Wideman, H. (2002). Factors contributing to teachers' successful implementation of IT. *Journal of Computer Assisted Learning*, 18, 480–488.
- [22] Harris, A. (2002). *School Improvement: What's in in for schools?* Routledge Falmer, New Fetter Lane, London
- [23] Holmes, B., Savage, T., & Tangney, B. (Eds). 2000. *Innovation in Learning in the Information Society: A Comparative International Study*. The Centre for Research in Information Technology in Education, Trinity College Dublin.
- [24] Hoopes, D.G & Postrel, S. (1999). Shared Knowledge: glitches, and Product Development Performance. *Strategic Management Journal*. 20, 837- 865.
- [25] Heron, R., and Hammond, F. (2001). Partnerships and educational benefits in post graduate nursing education. *Australasian Journal of Neuroscience*, 14(2), 18– 21.
- [26] Hetzner, S., Paulmann, E., & Rohleder, S. (2005). Experiences in initial and further teacher education in an elearning environment. Selected Papers from the 5th EDEN Open Classroom Conference: Addressing eLearning, Innovation Papers from the 5th EDEN Open Classroom Conference: Addressing eLearning, Innovation and Quality in Schooling. 5-7 October, Poitiers, France.
- [27] Jamaliah Abdul Hamid (2008). Knowledge strategies of school administrators and teachers. *International Journal of Educational Management*. 22 (3), 259 – 268.
- [28] Kathleen A.M (2000). Knowledge Management in the Twenty First Century: the Role of the Academic Librarian. Dissertation Abstract
- [29] Khe Foon Hew & Noriko Hara (2007a). Empirical study of motivators and barriers of teacher online knowledge sharing. *Education Technology Research Development*, 55, 573–595.
- [30] Khe Foon Hew & Noriko Hara (2007b). Knowledge sharing in online community health-care professionals. *Information Technology & People*, 20(3), 235 – 261.
- [31] Laszlo, A. & Laszlo, K., (2002). Evolving Knowledge for Development: The Role of Knowledge Management in a Changing World. *Journal of Knowledge Management*, 6(4), 400-412.
- [32] Lave, J., & Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press

- [33] Lee Tan Luck (2010) Supporting Community Learning with Digital Technologies and Knowledge for Societal Development. Paper presented at The Sixth Pan-Commonwealth Forum (PCF6) on Open Learning, Kochi, India.
- [34] Low H. H., Omain, S. Z. & Md Som, H. (2000), Knowledge sharing culture among small and medium size industries. Research report. Research Management Centre, Universiti Teknologi Malaysia.
- [35] McDermott, R. & O'Dell, C. (2001). Overcoming cultural barriers to sharing knowledge. *Journal of Knowledge Management*, 5(1), 76 - 85.
- [36] McFarlane, D (2008). Effectively Managing the 21st Century Knowledge Worker. *Journal of Knowledge Management Practice*, 9(1): Retrieved April 28, 2011 from <http://www.tlinc.com/article150.htm>.
- [37] Ministry of Higher Education (2006). Report by the committee to study, review and make recommendations concerning the development and direction of higher education in Malaysia: Towards excellence. Putrajaya; Malaysia
- [38] Ministry of Education Malaysia (2003). Ministry of Education Malaysia Annual Report. Putrajaya; Ministry of Education
- [39] Nor Azlinda Jaafar (2010). Principals' roles in supporting knowledge management activities in secondary schools. Unpublished Master of Education. Universiti Teknologi Malaysia.
- [40] Pan, S., & Scarbrough, H. (1999). Knowledge management in practice: an exploratory case study. *Technology Analysis & Strategic Management*, 11 (3), 359-374.
- [41] Parr, J., & Ward, L. (2006). Building on foundations: Creating an online community. *Journal of Technology and Teacher Education*, 14(4), 775-793.
- [42] Punesvary, M., Hazita Azman, Azizah Ya'acob & Siew Ming Thang (2010). Blogging in teacher professional development: Its role in building computer-assisted language teaching skills. *International Journal of Education and Development using Information and Communication Technology*, 6(3), 73 - 87.
- [43] Salbiah Ismail (2003). ICT in the Classroom: A Malaysian Perspective. Paper presented at the Asia and the Pacific Seminar -Workshop on Educational Technology, 2 – 9 September, Tokyo and Kyoto, Japan.
- [44] Schlager, M. S., & Fusco, J. (2003). Teacher professional development, technology, and communities of practice: Are we putting the cart before the horse? *The Information Society*, 19(3), 203–220.
- [45] Smith, R. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, 1(2), 199-218
- [46] Staples, A., Pugach, M.C., & Himes, D. (2005). Rethinking the technology integration challenge: Cases from three urban elementary schools. *Journal of Research on Technology in Education*, 37(3), 285-311.
- [47] Tyack, D., & Cuban, L. 1995. *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press
- [48] Tiwana, A. (2000). *The Knowledge Management Toolkit*. Upper Saddle River: Prentice Hall.
- [49] Wang, D. H., Yang, D., & Chai, J.K. (2000). *Knowledge management*. 1st. Ed. Beijing: Xue Yuan Publisher.
- [50] Bond, T & Fox., C (2007). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences*. New Jersey: Lawrence Erlbaum Associates Publishers.
- [51] Linacre, J.M., (2006) *Winsteps: Rasch Measurement Computer Program*. Version 3.60, Winsteps.com, Chicago.

Lokman Mohd Tahir, currently a senior lecturer in the Faculty of Education, Universiti Teknologi Malaysia. His area of interest is the study of knowledge management in education.

Mohd Fadzli Ali, currently a lecturer in the Department of Educational Technology and Multimedia, Universiti Teknologi Malaysia. His area of interest is e-learning in higher educational setting.

Tina Abdullah, currently a senior lecturer at the Language Academy, Universiti Teknologi Malaysia. Her area of interest is computers and language learning.

Mehmet Ozay, currently a senior lecturer at the Faculty of Education and his research area is sociology of education.

Haswadi Hassan, currently a lecturer in the Department of Computer Science in the Faculty of Computing, Universiti Teknologi Malaysia. His areas of interest are computer image processing and pattern recognition.

ADOPTION OF ICT-BASED SECONDARY EDUCATION FOR ACTUALIZING THE MILLENNIUM DEVELOPMENT GOALS IN NIGERIA

Odeniyi O.A¹, Ayinde A.Q², Lawal N.B³, Ghazaly-Agboola A.B⁴ and Kolade S.O⁵

1. Computer Science Department, Osun State College of Technology,
Esa-Oke, 234035/South West, Nigeria

2. Computer Science Department, Osun State College of Technology,
Esa-Oke, 234035/South West, Nigeria

3. Computer Science Department, Osun State College of Technology,
Esa-Oke, 23402/South West, Nigeria

4. Computer Science Department, Osun State College of Technology,
Esa-Oke, 234035/South West, Nigeria.

5. Computer Science and Engineering Department, LAUTECH,
Ogbomoso, 23402/South West, Nigeria.

Abstract

It is clear that the Nigeria government is making deliberate efforts toward achieving the Millennium Development Goals (MDGs). However this paper is interested in seeing the adoption of ICT-based secondary education for actualizing the MDGs in Nigeria. It touches upon the three dimensions of educational processes namely "learning", "teaching", and "educational environment" also focused on the action of learning.

Quality education can and should be sought everywhere, and not only in more affluent contexts. ICT integration into educational processes should be part of the effort towards improving them and making the most of available resources, even when these are meagre as in most developing countries like Nigeria. ICT also provides a glimmer of hope towards achieving the goal of quality universal education. The developing country has to find new methods that will dramatically improve both its children's schooling and its continuing education system. As in every other walk of modern life, the answer to the challenge of education for economic development will be positioned, in part, with technology.

Keywords: *Educational Environment, ICT, Learning, MDGs, Secondary Education*

1. Introduction

Nigeria as a nation cannot afford to be isolated from the wind of globalization. The wind of change is blowing rapidly across the globe. Changing economic, social and political situations in both developed and developing nations have combined to create needs for constant innovations and reforms in education.

Educational transformations are always the result and the symptom of social transformations in term of which they are to be explained. In order for people to feel at any particular moment in time the need to change its educational system, it is necessary that new ideas and needs have emerged in which the former system is no longer adequate.

In the developing countries, from the late 1950s to mid 1970s independence from colonial administrators, and in some cases new found wealth based on natural resources have contributed to a redefinition of social priorities and objectives.

For the fact that the society cannot do without reforms, the whole world came together to initiate the millennium development goals (MDGs). The MDGs are eight goals that the 191 United Nations member states have agreed to try to achieve by the year 2015. These goals are as listed below:

- (1) Eradicate extreme poverty and hunger;
- (2) Achieve universal primary education;
- (3) Promote gender equality and empower women;
- (4) Reduced child mortality;
- (5) Implore material health;
- (6) Combat HIV/AIDS, malaria, and other diseases;
- (7) Ensure environmental sustainability and
- (8) Develop a global partnership for development.

The goals as listed above portend that reforms are indispensable. In the same vein, reforms cannot do without education. The implication here is that whatever initiatives a government or an organisation evolves for improving its operations or society, must among others, be backed up with appropriate educational strategies, practise and activities/learning experiences. Also education is part of the struggle over power and hegemony in the society. This struggle manifests itself through curriculum content, through what constitutes official knowledge, adding that, the day-to-day interactions in the classroom are deeply ideological processes that shape people's ideas about the way the world works, and one's place in it.

In compliance with the MDGs, the Nigerian government produced a document, national economic empowerment and development strategy (NEEDS) representing a compendium of strategies for reforms in all sectors of her society. Under NEEDS, education (especially basic education) is considered the key bridge to the future. Education is so regarded as a key instrument to empower the children to take charge of their lives in the future. It equally targeted job creation, that is with the intention of drastically reducing the incidence of poverty and hence put people at the centre of development efforts. The document was also not silent on women. Its thrust in respect of women is to fully integrate them through enhancing their capacity to participate in the economic, social, political and cultural life of the country.

From the foregoing, it is clear that the government of Nigeria is making deliberate efforts toward achieving the MDGs. However this paper will be more interested in seeing the adoption of ICT-based secondary education for actualising the MDGs in Nigeria. The paper touches upon the three dimensions of educational processes, i.e. "learning", "teaching", and "educational environment", but focuses on the action of learning. It starts from the notion that all educational processes, particularly those said to be "student centred" should be towards good learning.

2. Literature Review

As Moore [4] points out, ICT also provides a glimmer of hope towards achieving the goal of quality universal education. The developing country has to find new methods that will dramatically improve both its children's schooling and its continuing education system. Guitert and Coderch [3] highlighted some of the new informational skills listed below, which offer particular value in present and future educational environments. Duarte [1] proposes a series of direct docent actions that while in the realm of *virtual learning environments* apply equally well to secondary education. Perhaps Resnick [7] exaggerates when affirming that "As new technologies continue to

quicken the pace of change in all parts of our lives, learning to become a better learner is far more important than learning to multiply fractions or memorizing the capitals of the world". Duarte [2] proposes that his models of learner-centred learning in universities could well be applied to secondary education. This includes greater freedom for the student/learner to use various types of supports, or the possibility to plan individual progress at the same time as regulating one's own working rhythm. According to Peters [5], teachers can take advantage at least four new types of possibilities based on ICTs. Certain conditions pointed out by Werry [8] in university settings, like de-formalization of working conditions or teachers' loss of control over the academic process, are even more evident in secondary education (this includes developing countries as well). Add to this low salaries and the growing level of conflict in the classroom (as is occurring in Spain, for example). Peters [5] underlines that Aristotle went as far as postulating that even the most abstract knowledge is base in sensory perception. According to Rashke (2002), learner-centric learning does not aim so much at mastering pre-established educational contents, but rather on the capacity to generate new contents.

3. Need for Adoption of ICT-Based Secondary Education in Nigeria.

Nigerian students' starts their secondary school education between the age of 9 to 14 years. We start from the assumption that in the secondary cycle it is possible to work on explicit pedagogical methods with the students. At this point in his/her education, the student possesses sufficient cognitive capacity about learning processes to enable teachers and students to build on educational practises introduced in earlier years. In other words this is the period for firming up learning/studying methods. Moreover, the student has attained sufficient social and intellectual capacity to consciously integrate the use of information and communication technologies (ICTs) particularly internet as a set of tools for new ways of learning.

The problems of the digital divide and other development divide must be mentioned at the outset. Clearly, the greater challenge is being able to educate all girls and boys in the world (Nigeria inclusive) which is moreover a well-accepted right of those children. The concepts hereby proposed imply prerequisites which are absent in too many cities and villages of the world (Nigeria inclusive), and therefore there are more directly applicable where basic educational standard had already been met.

If (ICTs) particularly internet changes educational environments along with the social context, it makes sense

that students will need to adapt to these changes. This implies that the student should “learn how to learn” in the new educational context. To be sure, knowing how to learn is always one of the key objectives pursued in any proper pedagogical space. But at the level of a given educational environment (and perhaps more widely in the entire educational system), we could characterise the main pedagogical challenge as figuring out institutionally how to best empower students to learn.

It is assumed that pedagogical techniques in the secondary cycle are fully developed and formalised. While in the primary cycle, the student has followed certain learning patterns. It is in high school when such practices can become explicit. The student is mature enough in cognitive terms to understand that besides the usual subjects (maths, literature, and history) she can deliberately apply various learning methods. Most of us remember our best high school teachers more for the methods they made us discover that for the specific content they taught us.

Extending to the tertiary cycle (university), we could argue that while retaining only a small amount of the content we learned earlier, the university taught us to think (which is another way of saying “taught us to learn”). It is about reasoning and methodologies, about ways to formulate and analyse a given problem, about the tools we use; all these knowledge we gained has shrunk with time. Indeed, from the perspective of learning and the educational environment, the time factor appears as a resource to be explicitly integrated in plans and methodologies, for example, in the context of a given classroom, questions should arise like “how much can one learn, in span of time ‘x’, about skills and knowledge related to the subject?”. The internet will help rationalize the time invested by public school teachers in the classroom. Technology tools can be included systematically in pedagogical models to make learning more efficient (less time for the same results), while maintaining or increasing its effectiveness (results in line with objectives). It was recorded that over seventy percent of Nigeria Universities had ICTs infrastructure fully implemented and from analysis it was observed that most year one students that finished from public secondary schools have challenges when writing a computer based examination or computer based mid semester test.

Less than five percent of the Nigeria public secondary schools have sustainable and effective ICTs infrastructure compared to the private secondary schools which are fully equipped with ICTs infrastructure. The public secondary schools are poorly funded by the government and this had adverse effect on both the students and the teachers in the e-learning world.

Questionnaire was distributed to ten different universities second year students in Nigeria on their educational

performance on the computer based test and examination they had in their first year at the university. Total of 10,000 questionnaires was distributed and 9880 were administered. The analysis show that 3927 students finished from a public school in Nigeria, 5093 students finished from a private secondary school, 804 students are from neighboring Africa countries and 56 students are from other continent.

Averagely, 12percent of the students from the public secondary school had GOOD grade, 36percent from the private secondary had GOOD grade, 71percent of the students from the neighboring Africa countries had GOOD grade and 89percent of the student from other continent had GOOD grade.

4. Conclusion

The adoption of ICT for secondary education is indispensable in actualizing the Nigeria’s education reform and achievement of the MDGs. It is in this regard ICT makes it possible for nations to meet development goals such as poverty reduction, basic healthcare and education. However, for ICT in education to promote the achievement of MDGS, there is need for the effective management and coordination of ICT resources. This is an area that demands a lot of contributions from school head teachers and especially schools principals. Therefore it is necessary for principals to understand clearly the meaning of ICT, it benefits to human beings especially in education and how ICT in education can be developed and implemented.

The following are the ideas which a Nigerian principal who is in the process of starting ICT in education should considered to ensure an e-learning environment for his/her students or teachers:

- (i) find out what is the current provisions in curriculum resources, for training and other needs?
- (ii) find out the school basic needs to be ensure a friendly e-learning environment for both the students and the teachers
- (iii) create an ICT section in the school development plan
- (iv) conduct a staff ICT training and must conduct weekly assessment for both the teachers and the students
- (v) review the curriculum
- (vi) consider approaches to teaching and the structuring of tasks
- (vii) consider how ICT can have impact on a range of learning approaches
- (viii) consider the ICT present learning environment and what is needed to deliver the curriculum
- (ix) determine how assessment, recording and reporting for ICT will be done
- (x) consider developing a policy statement for ICT across the school

(xi) determine how monitoring and evaluation of ICT will be carried out.

Below are the following indices that can be employed for a successful ICT implementation as a guide to principals:

- (i) motivated, challenged and required to think themselves;
- (ii) encouraged to be independent and confident users of technology;
- (iii) set tasks that encourage cooperation and collaboration to tackle them and solve given problems; and
- (iv) more responsive in the context of support for learning activities.

These indices of successful ICT in education should guide every principal during the development and implementation phases of ICT in education in his/her school. At the planning stage, they guide the selection of learning content and resources, while at the implementation stage they guide the process of evaluation. One under pinning requirement is that every school principal should endeavor to be ICT literate and to be competent in the use of technology. For sustainable ICT in education, the principal or administrator 'must have a broad understanding of the technical, curricular, administrative, financial and social dimensions of ICT use in education. Without this type of understanding, the principal cannot provide the leadership necessary for the achievement of the MDGs and actualizing education reforms through ICT-based secondary education.

It has been shown in this paper that ICT-based secondary education can lead to education reform and the achievement of the MDGs. However, this can only be possible when ICT is effectively implemented in schools and in the society. To promote the development and implementation of ICT-based secondary education, the following recommendations are made:

- (i) ICT education should be compulsory in all secondary schools in Nigeria.
- (ii) Serving teachers should be given the opportunity within a specific period to become ICT literate through in-service education
- (iii) Government should ensure the provision of basic ICT facilities in all secondary schools
- (iv) School principals should involve the parent
- (v) teacher associations in the provisions of ICT facilities in schools.
- (vi) Every secondary school should have an ICT coordinator.
- (vii) The state and federal ministries of education should ensure the provision of electricity in every secondary school.

References

- [1]Duart, Josep M.; Sangrà, Albert. (1999). "Formación universitaria por medio de la web :un modelo integrador para el aprendizaje superior". En: Duart, J.M.; Sangrà,A.*Aprenentatge i Virtualitat*(pp.23-49). Barcelona:Ediuc-Proa
- [2]Duart, Josep M (2003). Educar en valores en entornos virtuales de aprendizaje: realidades y mitos. UOC, <http://www.uoc.edu/dt/20173/index.html>
- [3]Guitert, Montse; Coderch, Jorge. (2001). "Como aprender y enseñar con Internet" *Cuadernos de Pedagogía*, 301 (pp. 56-63)
- [4] Moore, Michael C. (2002). "A personal view: Distance education, Development, and the problem of Culture in The Information Age". En: Venugopal Reddy, V; Manjulika, S.(ed.). *Towards virtualization: open distance learning*. (pp. 634-640). Kogan Page India, Pvt. Ltd
- [5] Peters, Otto (2000). "Digital Learning Environments: New Possibilities and Opportunities". *International Review of Research in Open and Distance Learning*. Vol. 1, no. 1
- [6] Raschke, Carl A. (2003) *The digital revolution and the Coming of the Postmodern University*. Routledge Farmer. London. Chapter 3, pp. 26–38
- [7] Resnick, Mitchel. (2002). "Rethinking Learning in the Digital Age". In *The Global Information Technology Report: Readiness for the Networked World*, edited by G. Kirkman. Oxford University Press. (pp. 32-37)
- [8] Werry, Chris (2001) "The work of Education in the Age of E-college". *First Monday, Peer Review Journal of the Internet*, vol. 6, no. 5, May 2001.

O.A Odeniyi obtained his B.Tech (Computer Science) from LAUTECH (1996). He held Post Graduate Diploma (PGD) in Education from National Teachers' Institute Kaduna (2006). He is presently a research student at the Department of Computer Science and Engineering, LAUTECH. He is a Lecturer 1 at Osun State College of Technology Esa Oke. He is a member of Computer Professionals Registration Council of Nigeria (CPN) and Nigeria Computer Society (NCS). His research areas are soft computing, ICT, Data mining and Database.

A.Q Ayinde obtained his B.Tech (Computer Science) from LAUTECH (2008). He is a research student in the Department of Computer Science and Engineering, Ladoko Akintola University of Technology (LAUTECH) and a Lecturer in the Department of Computer Science at Osun State College of Technology, Esa-Oke. His research areas include Data Mining, ICT, Soft Computing and Database. He is a member of the following

professional bodies Redhat Linux (RHCE Certified) Microsoft (MCITP Certified) and Information and Technology in Infrastructure Library (ITIL Certified).

N.B Lawal obtained his B.Tech (Computer Science) from FUTA (2001), He is a research student in the Department of Information Technology Management for Business, University of Greenwich, UK and he is a Lecturer II in the Department of Computer Science at Osun State College of Technology, Esa Oke He is a member of Computer Professionals Registration Council of Nigeria (CPN). His research areas are Soft Computing, ICT, Data mining and Artificial Intelligence.

A.B Ghazaly-Agbppla obtained his Higher National Diploma (Computer Science) from Federal Polytechnic Ede (2002).He held PGD in Education from the National Teachers' Institute Kaduna (2007).He is currently winding up his B.Tech (Computer Engineering) at LAUTECH. He is a Lecturer II in the Department of Computer Science at Osun State College of Technology, Esa Oke He is a member of Computer Professionals Registration Council of Nigeria (CPN) and Nigeria Computer Society. His research areas are Computer Networks, Modeling and ICT

S.O Kolade obtained his B.Tech (Computer Engineering) from LAUTECH (2008).He is a research student in the Department of Computer Science and Engineering, Ladoke Akintola University of Technology (LAUTECH).His research areas include Biometrics, ICT and Soft Computing.

Minimizing Handoff Latency and Packet Loss in NEMO

Fatimah Abdulnabi Salman¹, Emad Hassan Al-Hemairy²

¹ Networks Engineering Department, Al-Nahrain University
Baghdad, Iraq

² Internet Engineering Department, Al-Nahrain University
Baghdad, Iraq

Abstract

This paper proposes a handoff scheme for Network Mobility (NEMO) to minimize the handoff delay and packet loss. This scheme is acronymed as MMHM (Multiple mobile router handoff management) scheme that addresses the handoff management in multihomed mobile network. In the proposed scheme, the effect of router discovery and Duplicate Address Detection (DAD) that have direct impact on the handoff procedure is eliminated. The cooperation of multiple mobile routers in carrying the traffic of one another during handoff process results in minimum packet loss. The mathematical modeling is carried out proves that the proposed scheme reveals better performance compared with the NEMO basic support protocol.

Keywords: NEMO, Handoff, Multihomed, Multiple Mobile Routers.

1. Introduction

With the development of the various mobile communication technologies such as (laptops, smart phones and PDAs), the users expected to be always connected to the best available access network. They expect to benefit from the application and services on their mobile devices while they move. This integration of heterogeneous networks will, however, lead to heterogeneities in access technologies and network protocols. The need for providing continuous connectivity to the mobile users has brought the Internet Engineering Task Force (IETF) to develop the Mobile IPv6 that manages host mobility [1].

When several devices connected in a local area network or personal area network move together, MIPv6 would inefficiently handle and manage the mobility of the entire network and maintain the sessions for every moving device on the network. Mobility using MIPv6 would increase the signal overhead, power consumption, bandwidth consumption and manageability. Furthermore, MIPv6 protocol cannot handle network mobility. Therefore, the IETF created a working group named Network Mobility (NEMO) Working Group [2] with the objective of developing mechanisms that provide

permanent Internet connectivity to all Mobile Network Nodes (MNN) via their permanent IP addresses as well as maintain ongoing sessions as the mobile network changes its point of attachment to the Internet.

The NEMO working group has proposed NEMO basic support protocol [3] which is an extension of the MIPv6 that support network mobility management. They named a device which is a mobile router (MR) that provides continuous connections to the mobile network and manage the mobility of the mobile network rather than an individual mobile node. The mobile router is connected to the Internet through an egress interface and would act as the default gateway to the mobile network nodes. However, it is common that mobile devices may have more than one interface such as PDAs, smart phones, laptops, and may provide multiple connections to the Internet for the mobile device. In this case, the mobile device is called multihomed device. Many previous researches address the mobility management in multihoming environment to provide vigorous, ubiquitous Internet access [4-8]. A multihomed mobile network is a mobile network that is connected to the internet either via multiple network interfaces of the MR or via multiple mobile router. The multihomed mobile router (MMR) may equipped with different access technologies such as (UMTS, WLAN, GPRS and Bluetooth), thus it would perform vertical handoff from one access technology to another. On the other hand, the horizontal handoff is performed when the network moves within the same access technology.

The Multihoming provides constant access to the internet and improves the overall connectivity of the mobile network. Some benefits can be obtained from multihoming in NEMO like load balancing and sharing, fault tolerance, redundancy, and reliability. The configuration of multihomed NEMO [9] can be classified according to three variables x, y, z , where x, y, z refer to the number of multiple mobile routers exist in the mobile network, the number of HAs and the number of mobile network prefix respectively.

In this paper, the context (n,1,1) configuration of multihomed NEMO is used. The multiple mobile router handoff management (MMHM) scheme is proposed to provide continuous connectivity for the MNN and eliminate the handoff delay and packet loss during handoff. In this scheme, only one tunnel will be active to carry the traffic to/from the mobile network.

The remainder of this paper is organized as follows. Section 2 gives an overview on handoff in NEMO BSP. Section 3 presents some related work concerning multihoming in NEMO. Section 4 describes the handoff operation of the proposed MMRH scheme in detail. Section 5 presents the mathematical analysis and results. Finally section 5 concludes the paper.

2. Handoff in NEMO Basic Support

This section describes the basic operation of NEMO basic support protocol (NEMO BSP) with a brief description of the handoff procedure when a MR moves away from its home network.

2.1 NEMO Basic Support Operation

The IETF has proposed NEMO basic support protocol which is an extension of MIPv6 to manage the mobile network while moving from its home network and attached to a visited network. The NEMO BSP proposes one mobile router (MR) in a mobile network that works on behalf of all mobile nodes in performing mobility functions. The MR provides continuous connectivity to the MNNs within its network through an ingress interface. These MNNs are unaware of the network mobility and they don't perform any mobility functions as illustrated in Fig. 1.

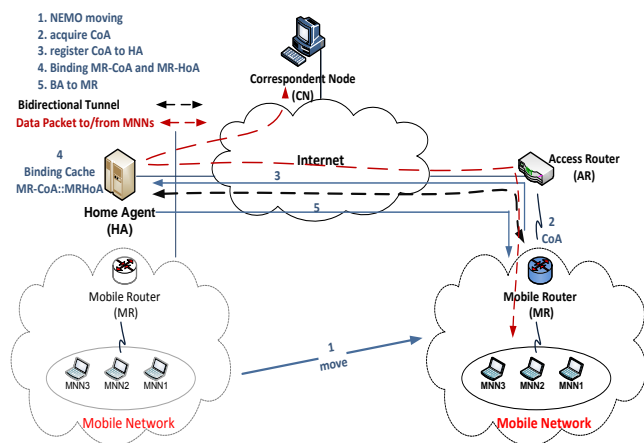


Fig. 1 NEMO BSP operation where numbers represent sequenced operation for NEMO

The MR in the mobile network has two addresses; the permanent address which is called the home of address (HoA) that is assigned to the MR in its home network. The temporary address is called Care of Address (CoA) that obtains at the visited network. The MNN within the mobile network obtains their addresses according to the MR prefix which is called the mobile network prefix (MNP) that remains fixed even if the mobile network moves to another point of attachment.

When the MR moves to another network (performs handoff) and acquires CoA, it sends binding update (BU) message to the home agent (HA) to register its CoA and bind it with MR-HoA. Once the MR receives binding acknowledgement (BA) from the HA, a bidirectional tunnel is established between the MR and HA. The end points of this bidirectional tunnel are the HoA and CoA of MR. All the outgoing and ingoing packets between the MNN inside the NEMO and a node outside the NEMO are intercepted by the HA and routed through this bidirectional tunnel.

2.2 Handoff Delay in NEMO

When a MR changes its point of attachment to the Internet, handoff occurs. The MR handoff goes through the following steps; each step causes a delay that affects the overall handoff delay:

- Link layer handoff (L2 handoff) delay: this delay represents the time when a MR changes its physical connection and associated to an access point. It is not necessary that MR changes its network if the access point is in the same network, in this situation the CoA doesn't change.
- Movement Detection (MD) delay: this delay is associated with the time interval of router discovery. It is the time that MR discovers that it has been disconnected from the HA and determines its network layer movement when it receives router advertisement (RA) messages from the new Access Router (AR) [10].
- DAD delay: this delay represents the time for verifying the uniqueness of CoA. The MR generates its CoA from the prefix contained in RA received from AR using IPv6 stateless autoconfiguration mechanism, and then verifies the address uniqueness using Duplicate Address Detection (DAD) process. The DAD procedure delay takes (1-2) sec [11].
- Registration delay: this delay represents the time of the BU/BA procedure of MR until the MR receives its first packet through its newest connection [11], [12].

It is obvious from the above handoff delay analysis, that the CoA configuration and DAD procedure have the great

impact for increasing the handoff delay that needs more attention.

3. Related Work

Many schemes have been proposed to address the handoff efficiency in multihomed NEMO. Some researchers have benefit from the multiple interfaces that equipped in the mobile router such as a mobile node with WLAN and GPRS. For example, Petander [11] addressed the handoff performance issue by proposing a novel Make Before Break (MBB) handoff scheme. In MBB scheme, the MR is equipped with two network interfaces, one for data communication and the other for scanning networks. These interfaces will take over the operation of each other, once a better connectivity is found; the scanning interface take over the data transmission and the other reverts to a scanning role. Chen [13] studied the Internet connectivity of multi-interfaced MR equipped with (WLAN-CDMA and GPRS) egress interfaces. An inter-interfaces handover decision algorithm is proposed to provide seamless handover between different interfaces.

Multiple mobile routers in the same mobile network have much attention to the researcher. Tsukada [14] proposed the multiple mobile router management (MMRM) scheme that suggested using multiple MRs that serve the MNNs. The primary MR (PMR) in NEMO can detect the new joining MR via the router advertisement broadcasted by the new MR. The PMR establishes MR-MR tunnel with the new MR and regard this tunnel as a virtual egress interface. Thus, the PMR is responsible for selecting a path for each node attached to the mobile network dynamically. This paper will focus on Tsukada work, with modification to network scenario using two mobile routers already exist in the mobile network instead of joining and leaving MR. Mobile routers will register their CoAs using their own HoAs and only one directional tunnel will be active at a time rather using multiple CoAs registration using the HoA of the PMR and using multiple directional tunnel for carrying the traffic. The handoff operation of the mobile network is demonstrated and its performance is analyzed concerning the handoff delay and packet loss problems. Parkash [15] proposed a seamless handover scheme across heterogeneous networks, by using multiple MRs, each MR has its own Home Agent (HA) that belongs to different Internet Service Providers (ISPs). They suggest that the MR that undergoing the handover process sent two binding update messages to its HA (the first one using HoA of the stayed MR instead of its current CoA, the second is to use the CoA of the MR that is not under handover process) to support Network Mobility Basic Support (NEMO BS).

Others used additional entity to manage the handoff or manage resources or traffic, such as using Intelligent Control Entity (ICE) and integrated into multihomed NEMO architecture as propose by Lin [16] to manage handover, resources and adaptation for void zone. The information collected from Access Routers (ARs) and MRs within ICE domain stored in AR_INF table and MR_INF table in ICE respectively, helped to choose the best MR and AR during handover process. Slimane [8] introduced Mobile Network Proxy (MNP) entity as a central gateway at the Mobile Network Level. It managed handoff and traffic distribution between the multiple network routers by using unidirectional HA-MR tunnel. Two principles components implemented into the MNP, the environment detector component to detect the changes occurred of MRs and the policy decision component for mobility and traffic management.

Other work solutions are based on higher layer extension protocols for addressing multihoming in mobile network. Multihomed SIP-NEMO proposed by Huang [17] to solve multiple egress gateways in a mobile network. In addition to the REGISTER and INVITE methods, the REFER, SUBSCRIBE and NOTIFY methods are further utilized to handle the multihoming concern. SIP-NMSs negotiate and synchronize each other using the SUBSCRIBE and NOTIFY messages. Abdul-Razzaq [18] proposed two techniques for HIP-NEMO to reduce the packet during handoff, the HIP Multi-Homing (HIPMH) technique and HIP NEMO Cell Switching (HIPNCS) technique. In the HIPMH, the MR is equipped with two egress interfaces to provide direct path and multiple routes to the Correspondent Node (CN). In HIPNCS, the modified NCS algorithm is implemented at MR to check the RA message broadcast by other routers (ARs or neighbor MRs). The handoff execution decision depends on the Internet Connectivity Strength (ICS) parameters.

4. Handoff Proposed Scheme

The goal of the proposed scheme is to achieve seamless handoff and reduce packet loss during handoff by redirecting the traffic of one MR to another in the same mobile network. The network model illustrated in Fig.2 is considered in our analysis and can be implemented in any vehicular network (bus, long vehicle and trains). It consists of two mobile routers (MR1 & MR2) that are connected through a wired link and registered to the same HA. Each MR maintains a cache to keep the state of binding on neighbor MRs which contains the fields (HoA, CoA, a one bit flag (T) for tunnel status (Open/closed), Lifetime). Multiple bidirectional tunnel will be established to the HA, but only one tunnel will be active at a time to carry the traffic for the MNNs in the mobile network. The opened

tunnel can be recognized through the flag (T) which is added to the BU message to indicate which tunnel is opened HA-MR1 tunnel or HA-MR2 tunnel.

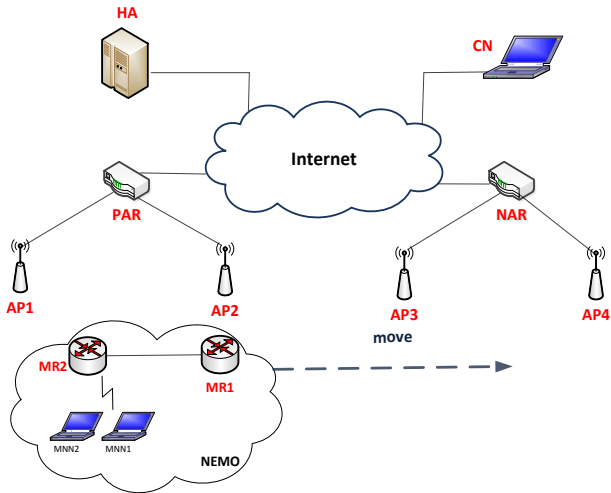


Fig. 2 Network diagram of the proposed scheme

It is assumed that the access routers (ARs) may have more than a couple of access points (APs). The AR collects the information about the APs belonging to it and stores it in a table named AP-table with one bit flag to indicate whether the AP is a boundary overlapping AP or not. The AP that has a common coverage area with the other AP that belongs to the other access router is called a boundary AP.

MR1 & MR2 are separated in different places, so that the handoff operation will happen at different times as the mobile network moves. MR1 scans its channel when it resides in the boundary overlapping area. When MR1 enters in a boundary area and detects two APs that belong to different AR; it must select the AP that belongs to the next AR (NAR) to perform handoff to prevent frequent handoff to the same subnet. Therefore, MR sends a query to its AR to determine whether the AP is a boundary AP or not. The AR checks the information maintained in the AP-table, and returns a message to MR1 to trigger to the AP belongs to the next AR.

The proposal assumes that the AR formulates the CoA for MR, does the DAD procedure and sends the unique CoA to MR through modified RA.

4.1 Care of Address Configuration and Duplicate Address Detection Procedure

Generally, MR configures its CoA according to the link layer information of MR and network prefix information of the AR using the IPv6 stateless Address Autoconfiguration mechanism. It is assumed that the AR will provide the CoAs to MR and performs the DAD

procedure by itself instead of MR. After validating the uniqueness of CoA by AR; it delivers the generated CoA to the MR through the modified RA.

MR1 requests for multiple modified RA one for itself and the other for MR2 through sending a modified RS message to AR. The basic RS message is modified by adding two flags, one bit flag (C) and two bits flag (R) as shown in Fig. 3. The C flag indicates that the MR sends the RS message requesting a new CoA from AR. The R flag indicates that MR is requesting more than one modified RA.

8 bits		8 bits		16 bits	
Type		Code		Checksum	
C	R	Reserved			
Source link-layer address options					

Fig. 3 Modified router solicitation message format

According to this modification, AR will respond with multiple modified RA messages. The modified RA message contains two flags, a single bit flag (C) to indicate that modified prefix information option includes the new CoA for a MR and the flag (R) to indicate that the multiple modified RA message has been sent. The modified RA message format is as shown in Fig. 4.

8 bits		8 bits				16 bits	
Type		Code				Checksum	
Cur.Hop Limit	M	O	H	C	R	Res	Router Lifetime
Reachable Time							
Retransmission Timer							
Prefix information option							

Fig. 4 Modified router advertisement message format

4.2 Handoff Operation

When the mobile network moves away from its home network, MR1 performs handoff first to transfers the packets destined to MNN through the next AR (NAR), it requests for new CoA for itself and for MR2 from the NAR, while MR2 still connected to previous AR (PAR) in order to perform lossless handoff. Figure 5 shows the handoff procedure of the proposed multihomed mobile network MMHM scheme to manage the mobility of the multihomed NEMO.

Both MR1 and MR2 are connected to PAR and MR2 is the default gateway for MNNs in the mobile network. When the mobile network moves and changes its point of attachment to the Internet, The following functions are executed.

MR1 detects that it's about to move away from its network. It sends the modified RS message with the C flag and R flag are set requesting for multiple modified RA messages that contain new CoAs from the NAR. The NAR will respond with two modified RA messages (RA1 & RA2) that have prefix information option that includes the complete and validate IP addresses (CoAs). MR1 will map this new CoA with the previous CoA of MR2 to establish a tunnel to transport the traffic for MR2. MR1 sends two BU messages to the HA and MR with T flag is set in order that the HA opens the tunnel with MR1 and closed the tunnel with MR2 to redirect the traffic to the NAR, then to MR1 and MR2. Also, MR2 will update its binding cache, and binds its previous CoA with the new CoA of MR1 with T flag is set to redirect the traffic coming from MNNs to MR1 then to NAR.

Then, MR1 sends the modified RA2 message to MR2 that contains the valid new CoA for MR2. When MR2 receives the new CoA from MR1, it concludes that it will perform handoff in the near future. So it scans the channel and performs handoff and register the new CoA with the HA by sending BU message with T flag sets to open the tunnel with MR2 and close the tunnel with MR1.

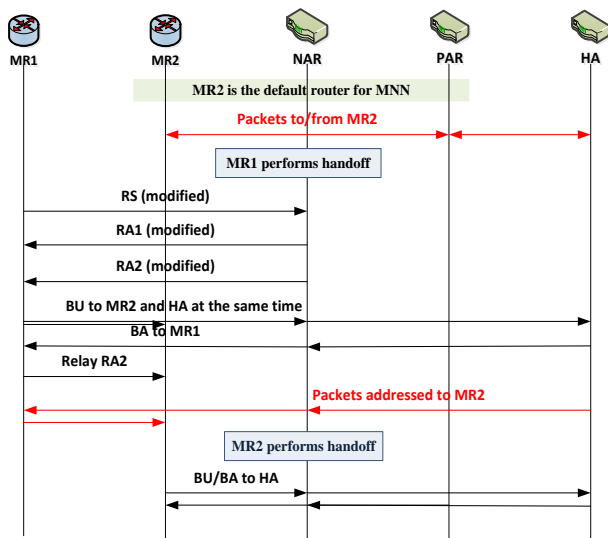


Fig. 5 Timing diagram for the proposed scheme

5. Performance Analysis

This section presents the assessment and comparison of the performance of the proposed scheme with the NEMO BSP. Handoff delay and packet loss are the critical metrics

used to evaluate the handoff mechanism. The default parameters values in table 1 are used. Generally, the delay of the packet transmission consists of processing delay, transmission delay and propagation delay. This can be denoted as Eq. (1):

$$\text{Packet Delay} = T_{proc} + T_{trans} + T_{prop} \quad (1)$$

Where T_{proc} is the time taken for processing a packet, T_{trans} is the time taken to transmit a packet as Eq. (2) and T_{prop} is the amount of time it takes for the signal to travel from the sender to the receiver, here it is referred as link latency.

$$\text{Transmission delay} = \frac{\text{packet size}}{\text{link bandwidth}} \quad (P/B \cdot W) \quad (2)$$

Let $t(p, d_{x-y})$ denote the transmission delay of a message of size P sent from 'x' to 'y'.

through the wireless link, it can be expressed as Eq. (3):

$$t(p, d_{x-y}) = T_{proc} + \frac{P}{B \cdot W_{wireless}} + L_{wireless} \quad (3)$$

through the wired link, it can be expressed as Eq. (4):

$$t(p, d_{x-y}) = \left(T_{proc} + \frac{P}{B \cdot W_{wire}} \right) \times d \quad (4)$$

Table-1 Simulation parameters values

Parameters	Symbols	Values
L2 handoff delay	T_{L2}	50ms
Wired link bandwidth	$B \cdot W_{wire}$	100Mbps
Wireless link bandwidth	$B \cdot W_{wireless}$	11Mbps
Wired link latency	L_{wire}	2ms
Wireless link latency	$L_{wireless}$	20ms
Packet arrival rate	λ_p	10 pkt/sec
Hops between MR and HA	d	2
Processing Delay	T_{proc}	2ms
Mobility speed	V	25m/sec
Packet size	P	512 byte
Binding Update size	P_{bu}	80 bytes
Binding Acknowledgement size	P_{ack}	40 bytes

5.1 Handoff Delay

The handoff delay is the interrupt time from the time MR leaves its network and the time that MR receives packets from another network.

The handoff delay of the NEMO BSP is composed of the layer 2 handoff delay (T_{L2}), the CoA configuration delay (T_{CoA}), the DAD procedure delay and the HA registration delay (T_{reg}) that is computed as Eq. (5).

$$T_{HO-NEMO\ BS} = T_{L2} + T_{CoA} + T_{DAD} + T_{reg} \quad (5)$$

T_{CoA} delay depends on the average of the minimum router interval and the maximum router interval (router discovery). The movement detection is Eq. (5).

$$(T_{md}) = \frac{1}{2} \left(\frac{MinRtrInterval + MaxRtrInterval}{2} \right) \quad (6)$$

T_{Reg} delay of sending BU message from MR to HA. This delay is equal to the sum of the delays of all the links delay between MR and HA, the delay between MR-AR (wireless link) and the delay between AR-HA (wired link) as shown in Eq. (7).

$$T_{reg} = \left(2T_{proc} + \frac{P_{bu} + P_{ack}}{B.W_{wireless}} + 2L_{wireless} \right) + \left(2T_{proc} + \frac{P_{bu} + P_{ack}}{B.W_{wire}} + 2L_{wire} \right) \times d \quad (7)$$

The handoff delay of the MMHM scheme, for MR1 is composed of layer2 handoff delay, the RA1 delay (T_{RA1}), the RA2 delay (T_{RA2}) and the HA registration delay (T_{reg}) as Eq. (8). For MR2 handoff delay as Eq. (9) is composed of only the layer2 handoff delay T_{L2} and HA registration delay T_{reg} .

$$T_{HO-MMHM(MR1)} = T_{L2} + T_{RA1} + T_{RA2} + T_{reg} \quad (8)$$

$$T_{HO-MMHM(MR2)} = T_{L2} + T_{reg} \quad (9)$$

Figure 6 shows the impact of the number of hops on handoff delay. For NEMO support and the proposed scheme, the handoff delay increases with the increase of the hops between MR and HA. It can be noticed that the proposed performance is better than the Basic NEMO.

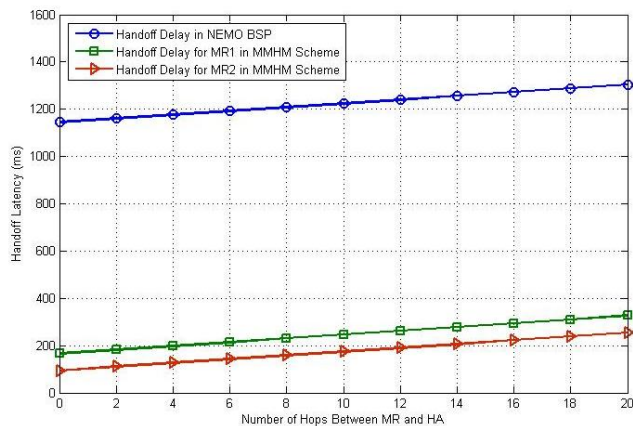


Fig. 6 Handoff delay vs number of hops between MR and HA

Figure 7 illustrate the handoff delay with respect to the router discovery. For the proposed scheme, only MR1 is affected by the router discovery since it performs handoff first and need to perform router discovery, MR2 is not affected by it since it relies on MR1 to deliver it the CoA that acquired from NAR.

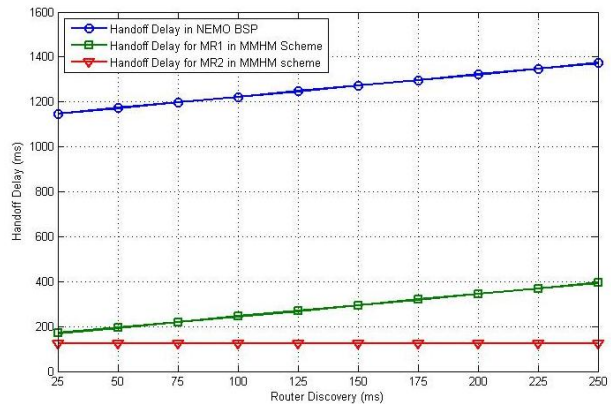


Fig. 7 Impact of Router Discovery on Handoff delay

5.2 Packet Loss

It represents the dropped packet during handoff; it mainly depends on the handoff delay. The packet loss gives an indication of the effect of handoff in the application. The packet loss increases during handoff with the increase in data rate. Here the packet loss is measured according to the offered data rate and as function of packet arrival rate.

Starting with an offered data rate of 100 Kbps and ending with 1000 Kbps by steps of 100 Kbits, the packet loss is as shown in figure 8. It can be noticed that the MMHM perform more efficient than NEMO BSP which means that the MMHM is more suitable for real time applications that send with high rate.

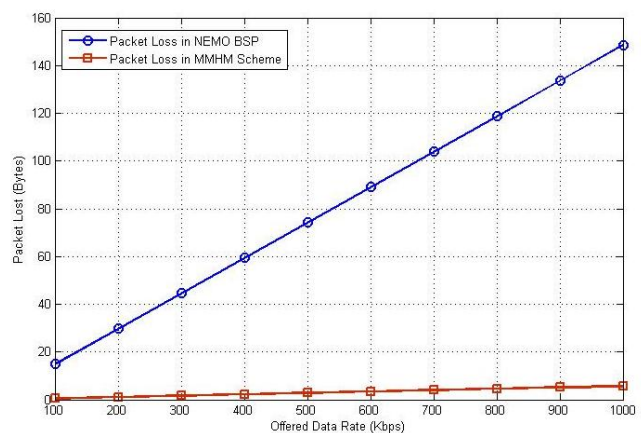


Fig. 8 Packet Loss as a function of offered data rate

Figure 9 shows the packet loss as a function of packet arrival rate (λ_p) as Eq.(10). It is obvious from the results that the proposed MMHM scheme gives better performance than the NEMO when the number of packet arrival of MNN is increased.

$$PL = N \times \lambda_p \times T_{HO} \quad (10)$$

N: the number of MNN having an active communication session through MR which is 20.

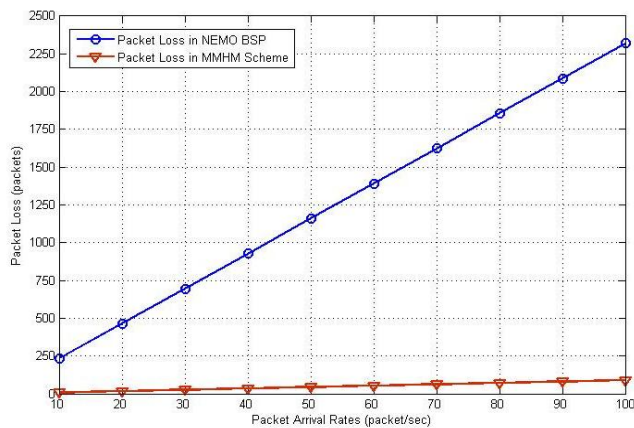


Fig. 9 Packet Loss as a function of packet arrival rate

6. Conclusion

In this paper, a multiple mobile router handoff management (MMHM) scheme is proposed to provide seamless handoff in multihomed mobile network. The multihomed NEMO configuration of the context (n,1,1) is used in the work. In this scheme, more than two mobile routers may exist in the mobile network that spatially separated from each other. These multiple mobile routers cooperate in delivering the packets destined to the MNNs in the mobile network that provide no interruption for services time, and reduce packet loss during handoff. The proposed scheme adds an overhead for tunneling the packet between MRs, but the effect of this tunnel is not remarkable compared with the results of handoff delay and packet loss. By avoiding the router discovery, DAD procedure and analyzing the results of the proposed MMHM scheme and NEMO BSP, it is clearly noticeable that the MMHM scheme performs more efficient than NEMO BSP for real time applications.

References

- [1] D. Johnson, C. Perkins, and J. Arkko, "Mobility Support in IPv6", RFC 3775, IETF, 2004.
- [2] IETF Network Mobility (NEMO) Working Group, <http://www.ietf.org/wg/nemo/charter>.
- [3] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert, "Network Mobility (nemo) basic support protocol", RFC 3963, IETF, 2005.
- [4] C. Ng, and T. Ernst, "Multiple Access Interfaces for Mobile Nodes and Networks", in 12th International Conference on Networks (ICON), 2004.

- [5] L. Suci, J. Bonnin, K. Guilloard, and T. Ernst, "Multiple Network Interfaces Management for Mobile Routers", in 5th International Conference on ITS Telecommunications (ITST), 2005.
- [6] L. Li-Hua, L. Yuan-An, and W. Qiu-Tian, "Mobility Management for Multihomed Mobile Networks using Backup Registration", in 2nd International Conference on Internet Multimedia Services Architecture and Applications (IMSAA), 2008.
- [7] C. Park, N. Choi, E. Paik, T. Kwon, and Y. Choi, "Multiple Interface/Prefix Selection for Virtual Mobile Networks", in 10th International Conference On Advanced Communication Technology, 2008.
- [8] Z. Slimane, M. Feham, and A. Abdelmalek, "A Seamless and Transparent MN-Proxy based Mobility Support For (n,n,1) Multihomed NEMO Model", International Journal of Computer Science and Network Security (IJCSNS), Vol.10, No.4, April 2010, pp.306-313.
- [9] C. Ng, E. Paik, T. Ernst, and M. Bagnulo. "Analysis of Multihoming in Network Mobility Support", RFC4980, IETF, 2007.
- [10] S. Islam, A. Hassan, R. A. Saeed, O. O. Khalifa, M. K. Hasan, O. Mahmoud, J. Daoud, and S. A. Hameed, "Mobility Management Schemes in NEMO to Achieve Seamless Handoff: A Qualitative and Quantitative Analysis", Australian Journal of Basic and Applied Sciences, 2011, Vol.5, N.60, pp.390-402.
- [11] H. Petander, E. Perera, K. C. Lan, and A. Seneviratne, "Measuring and Improving the Performance of Network Mobility Management in IPv6 Networks", IEEE Journal on selected Areas in Communications, 2006.
- [12] Z. Slimane, M. Feham, and A. Abdelmalek, "Seamless Infrastructure independent Multi Homed NEMO Handoff Using Effective and Timely IEE 802.21 MIH triggers", International Journal of Wireless & Mobile Networks (IJWMN), 2012, Vol. 4, No.3.
- [13] X. Chen, H. Zhang, Y. Chang, and H. Chao, "Experimentation and Performance Analysis of Multi-interface Mobile Router Scheme", Simulation Modeling Practice and Theory, 2010, pp. 407-415.
- [14] M. Tsukada, T. Ernst, R. Wakikawa, and K. Mitsuya, "Dynamic Management of Multiple Mobile Routers", in 13th IEEE International Conference on Networks, November 2005, vol. 2, pp. 1108-1113.
- [15] A. Prakash, R. Verma, R. Tripathi, and K. Naik, "Multiple Mobile Routers based Seamless Handover Scheme for Next Generation Heterogeneous Networks", in 1st International Conference on Network & Communication, IEEE, 2009, pp.72-77.
- [16] H. Lin, and H. Laboid, "Hybrid Handover Optimization for Multiple Mobile Routers-based Multihomed NEMO Networks", in IEEE international Conference on Pervasive Services, July 2007, pp.136-144.
- [17] M. Huang, C. H. Leet, and P. Tseng, "Multiple Router Management for SIP-based Network Mobility", IEEE, 2009, pp.863-868.
- [18] H. A. Abdul-Razzaq, "Network Mobility Management for Efficient Routing", M.S Thesis, Information and Communication and Engineering Department, Al-Nahrain University, Baghdad, Iraq, 2011.

Fatimah Abdulnabi has a B.Sc. in Computer Engineering, now she is M.Sc. student in Networks Engineering at Al-Nahrain University, Baghdad – Iraq.

Emad H. Al-Hemiary has a B.Sc. (1993), M.Sc. (1996) and Ph.D. (2001) in Electronics and Communications Engineering from Al-Nahrain University, Baghdad – Iraq. He is currently the head of Department of Internet Engineering in the College of Information Engineering at Al-Nahrain University. His fields of Interests includes Wireless Communications, Computer Networks and Next generation Networks. He is a member of the Iraqi Telecommunications Research Center and has many published papers in his fields of interest.

A PROPOSED PERMUTATION SCHEME BASED ON 3-D CHAOTIC SYSTEM FOR ENCRYPTING THE COLORED IMAGES

Osama M. Abu Zaid¹, Nawal A. El-Fishawy², and E. M. Nigm³

¹ Ph. D. Researcher in Computer Science, Department of Mathematics, Faculty of Sciences, Zagazig University, Egypt.

² Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Egypt.

³ Department of Mathematics, Faculty of Sciences, Zagazig University, Egypt.

Abstract

In this paper, a proposed permutation scheme based on three dimension chaotic map system will be presented. Chen's chaotic system is 3-D chaotic map system, which will be used to obtain the proposed permutation scheme. A proposed permutation scheme will be designated as PPS3DCS. It will be applied on two different color's frequencies colored-images. A proposed scheme (PPS3DCS), which Contains permutation procedure based on Chen's chaotic system is used to shuffle the positions of pixels of the colored plain-image. PPS3DCS will be applied on all color's channels of the image; Red, Green, and Blue. The expectant results of several experiments, statistical analysis, key sensitivity tests, and information entropy analysis will show that the proposed permutation scheme (PPS3DCS) is a good encryption scheme to provides an efficient and secure way for confusing or encrypting the colored images.

Keywords chaotic map; permutation; Image encryption; 3-D Chaotic map system.

1. Introduction

This age of communications revolution which necessitates multimedia transmission in a secure manner. encryption is important in transferring image through the communication networks to protect it against reading, alteration of its content, adding false information, or deleting part of its content.

Image encryption schemes have been increasingly studied to meet the demand for real-time secure image transmission over the networks.

Chaotic maps are very complicated nonlinear dynamic systems, which are applied in the field of figure correspondence and encryption [1-3], because they are very sensitive to initial conditions and can generate good pseudorandom sequences.

Chaotic systems have many important properties, such as the sensitive dependence on initial conditions and system parameters, pseudorandom property, non-periodicity and topological transitivity, etc. Most properties meet some requirements such as diffusion and mixing in the sense of cryptography[4]. Therefore, chaotic cryptosystems have more useful and practical applications.

Recently, a number of chaos-based encryption schemes have been proposed. Some of them are based on one-dimensional chaotic maps and are applied to data sequence or document encryption [5,6]. For image encryption, two-dimensional (2D) or higher-dimensional chaotic maps are naturally employed as the image can be considered as a 2D array of pixels [7-9]. The colored image consist of three 2D arrays of pixels for the color channels R, G, and B.

This paper will introduce a proposed permutation scheme for encrypting the colored images based on the 3-D chaotic map system (Chen's chaotic system). A proposed permutation scheme will be designated in this paper as (PPS3DCS). The permutation procedure based on Chen 's chaotic system is used to shuffle the positions of pixels of the colored plain-image.

The proposed permutation scheme PPS3DCS will be applied on all color's channels of the image; Red, Green, and Blue.

This paper is organized as follows. Section 2, will presents an overview on Chen's chaotic map system. Section 3, and its subsections 3.1, will discuss the proposed permutation scheme (PPS3DCS). Section 4, and its subsections 4.1 and 4.2 will present the experimental results and analysis by implementing statistical analysis and security analysis tests. Section 5, will discuss the final conclusion.

2. An Overview on Chen's Chaotic Systems

In this section, an overview on Chen's chaotic map system as important one of the three dimension chaotic map systems, which is used in this work.

Chen's chaotic map system is described by formula 1 which illustrates a set of the three differential equations of Chen's chaotic map system. [10-13]

$$\begin{cases} x = a(y_0 - x_0) \\ y = (c - a)x_0 - x_0z_0 + cy_0 \\ z = x_0y_0 - bz_0 \end{cases} \quad (1)$$

where $a > 0$, $b > 0$ and c such that $(2c > a)$ are parameters of the system [14]. Chen's system is chaotic when the parameters have the values; $a = 35$, $b = 3$ and $c \in [20, 28.4]$.

When $a = 35$, $b = 3$, and $c = 28$; it has a chaotic attractor as shown in Fig.1. It has been experienced that Chen's chaotic system is relatively difficult due to the prominent three-dimensional and complex dynamic property[10]. Recently, the study about Chen's chaotic map system has attracted many researchers' attention.

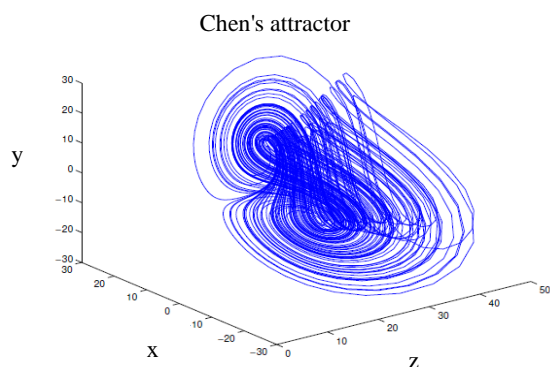


Fig. 1 Chaotic behavior of Chen's system

A very good performance for Chen's chaotic map at the parameters $a = 35$, $b = 3$, $c = 28$, the initial values $x_0 = 0$, $y_0 = 1$, $z_0 = 0$, and $h = 0.055555$ such that h is the step of the sequence [10].

3. A Proposed Scheme (PPS3DCS)

In this section, the proposed permutation scheme (PPS3DCS) based on Chen's chaotic systems is presented. The proposed scheme (PPS3DCS) consists of the permutation (confusion) procedure and the re-permutation (re-confusion) procedure. In this part of the paper the permutation (confusion) procedure only is designed and discussed because The re-confusion (decryption) procedure is the reversed technique of the permutation procedure.

3.1 The Permutation Procedure

The proposed permutation scheme (PPS3DCS) is designed to permute the positions of the pixels of the image, i.e. shuffling the positions of pixels of the image.

The proposed permutation scheme (PPS3DCS) consists of five steps of operations as following:

Step1: Obtain the $a1$, $a2$ and $a3$ matrixes (the three color components Red, Green and Blue) of the color image of size $m \times n \times 3$, respectively. $a1$ represents $m \times n$ matrix for the red, $a2$ represents $m \times n$ matrix for the green, and $a3$ represents $m \times n$ matrix for the blue. Afterwards, each color's matrix (including $a1$, $a2$ and $a3$) is reshaped by Matlab into one dimension matrix (vector) of integers within $\{0, 1 \dots 255\}$, wherein length of the vector is $si = m \times n$. Then, the so obtained three vectors ($aa1$, $aa2$, and $aa3$) represent the plaintext which will be permuted and confused.

Step2: Obtain the sequences XX , YY , and ZZ (1-D matrixes) as in formula 2 which are generated by Chen's chaotic system at $a = 35$, $b = 3$, $c = 28$, the initial values $x_0 = 0+k$, $y_0 = 1+k$, $z_0 = 0+k$, and $h = 0.055555$.

$$\begin{aligned} XX(i) &= \text{floor}(x) \text{ MOD } 256; \\ YY(i) &= \text{floor}(y) \text{ MOD } 256; \\ ZZ(i) &= \text{floor}(z) \text{ MOD } 256; \end{aligned} \quad (2)$$

Where i is from 1 to si . Values of x , y , and z are obtained from the three equations of Chen's chaotic system in formula 1. k is obtained by formula 3, where it is used to modify the keys in the proposed algorithm.

$$k = (k1+k2+k3)/10^{13} \quad (3)$$

Formula 4 is employed to generate the values of $k1$, $k2$ and $k3$ which are used to obtain k .

$$\begin{aligned} k1 &= \sum_{i=1}^m \sum_{j=1}^n a1(i, j) \\ k2 &= \sum_{i=1}^m \sum_{j=1}^n a2(i, j) \\ k3 &= \sum_{i=1}^m \sum_{j=1}^n a3(i, j) \end{aligned} \quad (4)$$

Step3: The matrixes XX , YY , and ZZ are sorted in descending sort by using Matlab function (sort). The Matrixes $XX1$, $YY1$, and $ZZ1$ are produced from sorting of the matrixes XX , YY , and ZZ respectively.

For example, let suppose $XX=[125 \ 3 \ 4 \ 10 \ 9 \ 5 \ 20 \ 8 \ 155 \ 255]$, after apply the function of descending sort; the result is $XX1=[255 \ 155 \ 125 \ 20 \ 10 \ 9 \ 8 \ 5 \ 4 \ 3]$. In position expression; the positions $[1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10]$ shifted to the positions $[3 \ 10 \ 9 \ 5 \ 6 \ 8 \ 4 \ 7 \ 2 \ 1]$.

Step4: The reshaped matrixes $aa1$, $aa2$ and $aa3$ are rearranged respectively according to the position of XX in $XX1$, the position of YY in $YY1$, and the position of ZZ in $ZZ1$. The sequences ar , ag , and ab which are obtained

from rearranging process of *aa1*, *aa2*, and *aa3* respectively.

For example, let suppose *aa1*=[125 56 90 42 50 220 120 255 65 35], according to the position of *XX* in *XX1* as in example of step3; the result is *ar* = [35 65 125 120 42 50 255 220 90 56].

Step5: Obtain the matrixes *arp*, *agp*, and *abp* (the permuted matrixes of the color's matrixes *a1*, *a2*, and *a3*), which are produced respectively by reshaping the sequences *ar*, *ag*, and *ab* from one dimension to the matrixes of two dimension $m \times n$.

According to the permutation (confusion) scheme, the position of any pixel in *a1*, *a2*, or *a3* is different with its position in *arp*, *agp*, or *abp* respectively, which will lead to be strong for the attacks.

4. Experimental Results and Analysis

In this paper, a practical programs of a proposed permutation scheme (PPS3DCS) and a practical programs of all experimental and security analysis tests are designed by MATLAB 7.0 on windows 7 system on Laptop computer with Intel CORE I₃ Processor, 3.0 GB RAM. All programs have been applied on two different color's frequencies colored-image (*flower.bmp* and *pepper.bmp*) as a plain-images of the size 120×120 pixels, which are shown in Fig. 2(a) and Fig. 3(a) respectively.

4.1 Statistical Analysis

To examine the quality of encryption and the stability via statistical attacks, the histogram is calculated for all color's channels R, G, B of the plain-images, correlation coefficient (CC) between each of color's channels R, G, B of the plain-image and the corresponding channels of the permuted-image, the correlation analysis of two adjacent pixels with the directions horizontal (HC), and vertical (VC) for all color's channels *R*, *G*, *B* of the encrypted-image.

4.1.1 Histogram Analysis

The plain colored-images (*flower.bmp* and *pepper.bmp*) of the size 120×120 pixels are shown in Fig.2(a) and Fig.3(a) respectively, and the histogram for *R*, *G*, *B* of these images is shown in Fig.2(b, c, d) and Fig.3(b, c, d) respectively.

Figure 4(a) and Fig.5(a) show the shuffled-images for *flower.bmp* and *pepper.bmp* which are produced from applying the proposed permutation scheme (PPS3DCS). The histogram for *R*, *G*, *B* of these images is shown in Fig.4(b, c, d) and Fig.5(b, c, d) respectively.

Figures 4 and 5 show that the histograms of the permuted (shuffled)-images are the same of the plain-images.

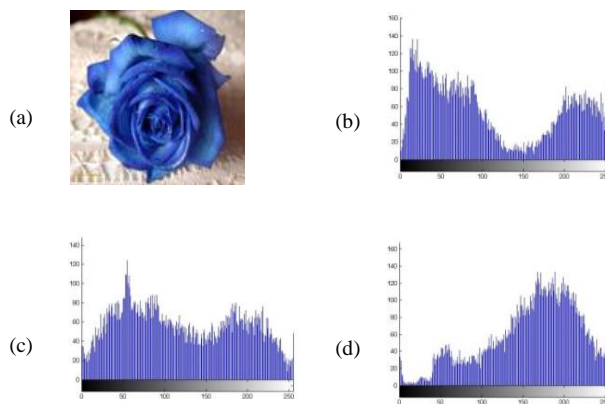


Fig. 2 The first plain-image and its histogram: (a) the image (*flower.bmp*); (b) histogram of R; (c) histogram of G; (d) histogram of B.

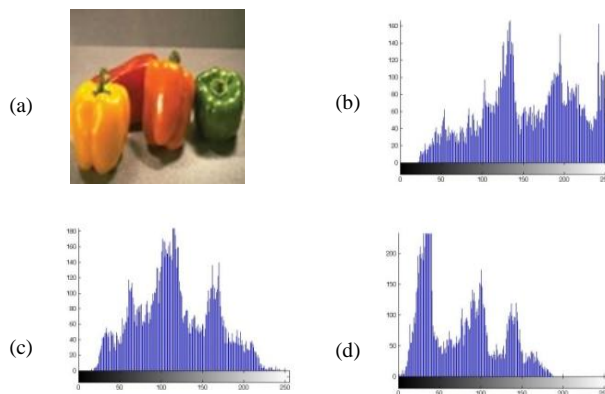


Fig. 3 The second plain-image and its histogram:(a) the image (*pepper.bmp*); (b) histogram of R; (c) histogram of G; (d) histogram of B.

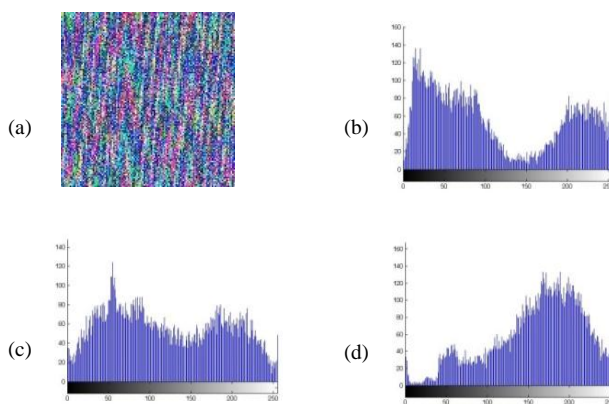


Fig. 4 The shuffled-image for *flower.bmp* and its histogram: (a) the shuffled-image; (b) histogram of R; (c) histogram of G; (d) histogram of B.

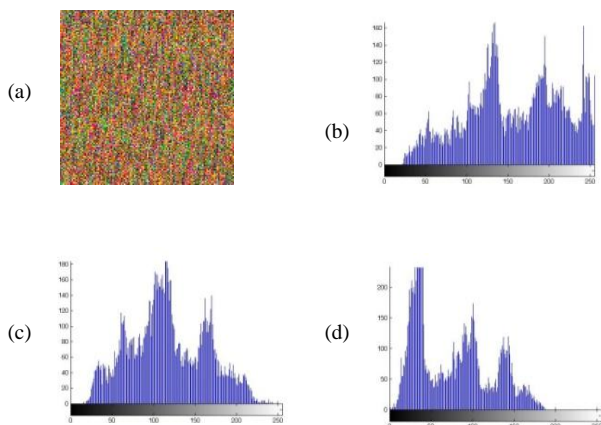


Fig. 5 The shuffled-image for *pepper.bmp* and its histogram: (a) the shuffled-image; (b) histogram of R; (c) histogram of G; (d) histogram of B.

From all previous figures of permuted -images and its histograms, as anyone can see, The proposed permutation scheme (PPS3DCS) is a complicated and very good procedure for disguise any countenance of the image without changing its histogram. Also, anyone can observe, the proposed scheme (PPS3DCS) is qualification for encrypting both the low frequencies colored-image (*flower.bmp*) and the high frequencies colored-image (*pepper.bmp*).

4.1.2 Correlation Coefficient Analysis

The correlation coefficient equals one if they are highly dependent, i.e. the encryption process failed in hiding the details of the plain-image. If the correlation coefficient equals zero, then the plain-image and its encryption are totally different. So, success of the encryption process means smaller values of the CC [15]. The CC is measured by formula 5:

$$CC = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^N (x_i - E(x))(y_i - E(y))}{\sqrt{\sum_{i=1}^N (x_i - E(x))^2} \sqrt{\sum_{i=1}^N (y_i - E(y))^2}} \quad (5)$$

$$\text{where } E(x) = \frac{1}{N} \sum_{i=1}^N x_i$$

where x and y are gray-scale pixel values of the plain and encrypted images. The CC is measured for each color's channel (R, G, B) of any colored-image.

Table 1: Results of CC analysis for encrypting *flower.bmp* and *pepper.bmp* by PPS3DCS.

Modes	CC analysis results		
	R	G	B
<i>Flower.bmp</i>	0.0127	-0.0113	0.0160
<i>pepper.bmp</i>	-0.0037	-0.00038	-0.0029

Table 1 and Fig. 6, illustrate that the proposed permutation scheme (PPS3DCS) achieves small values (very far from one and near to zero) of CC for the two images, so a PPS3DCS is a complicated and a good algorithm for encrypting the images. Also, the results of CC is better with the high frequencies colors image than the other image.

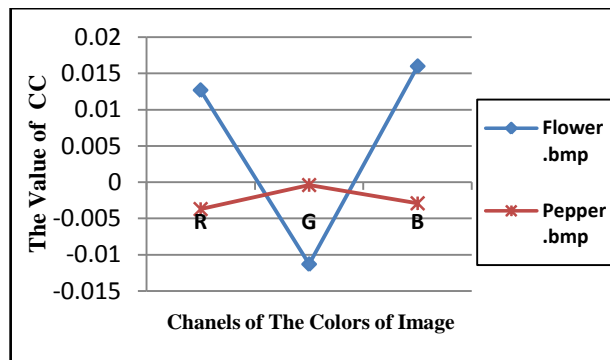


Fig. 6 Values of CC analysis for permuted images of *flower.bmp* and *pepper.bmp*

4.1.3 Correlation Analysis of Two Adjacent Pixels

It is well known that the adjacent pixels of an image have very high correlation coefficients in horizontal and vertical directions. The following formulas is employed to test the correlation between two horizontally adjacent pixels (designed as **HC**) and two vertically adjacent pixels (designed as **VC**) respectively, in plain images and permuted images, the following procedure was carried out. First, select 900 pairs of two adjacent pixels from an image. Then, calculate the correlation coefficient r_{xy} of each pair by using the following formulas [10,11]:

$$E(x) = \frac{1}{N} \sum_{i=1}^N x_i, D(x) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2 \quad (6)$$

$$cov(x, y) = E(x - E(x)) (y - E(y)) \quad (7)$$

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)} \sqrt{D(y)}} \quad (8)$$

Where x and y denote two adjacent pixels, and N is the total number of duplets (x, y) obtained from the image.

Table 2 illustrates the results of HC and VC analysis for the two plain colored-images.

Table 3, Fig. 7, and Fig. 8 illustrate the results of HC and VC analysis for the two permuted-images, which are produced by applying the proposed permutation scheme (PPS3DCS) on the plain-images.

Table 2: Results of HC and VC analysis for the plain images *flower.bmp* and *pepper.bmp*.

	<i>(flower.bmp)</i>			<i>(pepper.bmp)</i>		
	R	G	B	R	G	B
HC	0.9664	0.9670	0.9749	0.9980	0.9968	0.9964
VC	0.9709	0.9613	0.9479	0.9822	0.9739	0.9772

Table 3: Results of HC and VC analysis for the permuted images of *flower.bmp* and *pepper.bmp* by applying the PPS3DCS.

	<i>(flower.bmp)</i>			<i>(pepper.bmp)</i>		
	R	G	B	R	G	B
HC	0.1069	-0.0091	-0.0032	-0.0214	0.0041	-0.0010
VC	0.2798	0.3087	0.1598	0.0725	0.0254	0.0549

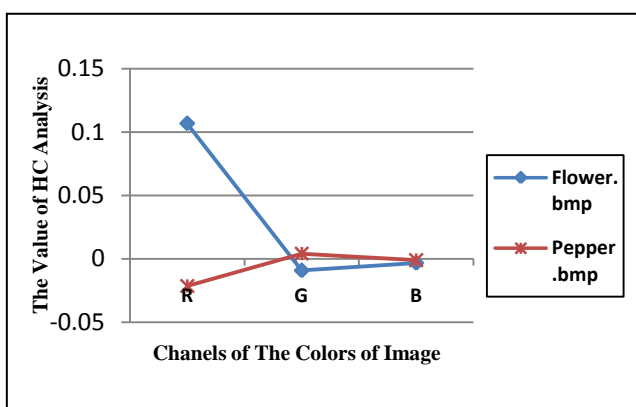


Fig. 7 Values of HC analysis for permuted images of *flower.bmp* and *pepper.bmp*

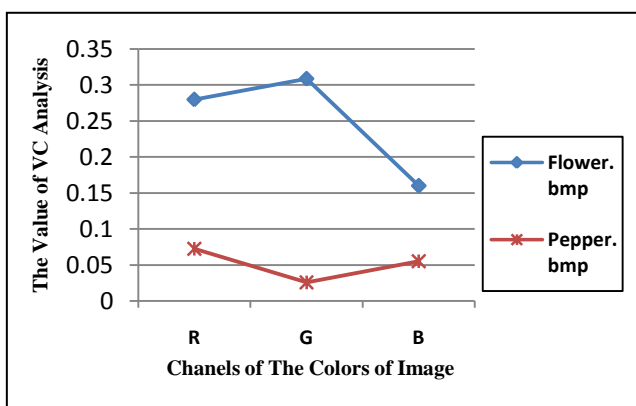


Fig. 8 Values of VC analysis for permuted images of *flower.bmp* and *pepper.bmp*

According to Table 2, anyone can observe, the results of HC and VC for the correlation analysis of two adjacent pixels for both the two plain-images are approach to 1, implying that high correlation exists among pixels.

According to Table 3, Fig. 7, and Fig. 8, the results of HC and VC for the correlation analysis of two adjacent pixels for both the two permuted-images with the modes are approach to 0, implying that no detectable correlation exists among pixels. Therefore the proposed permutation scheme (PPS3DCS) can protect the permuted-image from statistical attacks.

Also, According to Table 3, Fig. 7, and Fig. 8, the results of HC and VC is better with the high frequencies colors image than the other image.

4.2 Security Analysis

A good encryption algorithm should resist most kinds of known attacks, also it must be achieves sensitive to any little change in secret keys and a good values for the information entropy analysis.

In the proposed permutation scheme (PPS3DCS), the parameters *a*, *b*, *c*, and *h*, the initial values x_0 , y_0 , and z_0 are used as a secret keys.

4.2.1 The Key Sensitivity Analysis

The experimental results demonstrate that the proposed scheme (PPS3DCS) is very sensitive to the secret keys mismatch. The decrypted images by using PPS3DCS are the same of the original images, where are decrypted by using PPS3DCS with $a=35$, $b=3$, $c=28$, $h=0.055555$, $x_0=0+k$, $y_0=1+k$, and $z_0=0+k$ to produce the original image.

The experimental results for applying PPS3DCS on *pepper.bmp* demonstrate that the proposed scheme (PPS3DCS) is very sensitive to the secret keys *a* mismatch (10^{-14}), *b* mismatch (10^{-15}), *c* mismatch (10^{-14}), *h* mismatch (10^{-16}), x_0 mismatch (10^{-16}), y_0 mismatch (10^{-15}), and z_0 mismatch (10^{-14}).

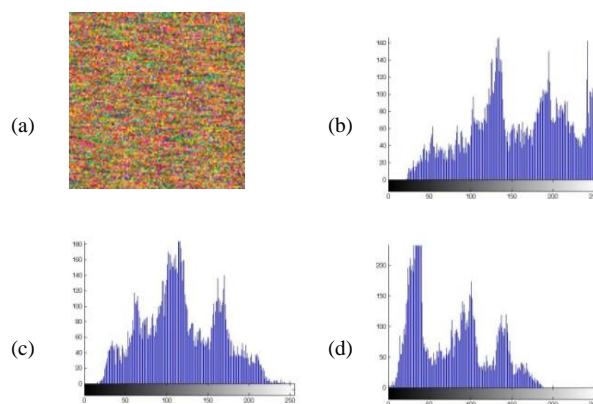


Fig. 9 The sensitivity to the secret key *b* of PPS3DCS for decrypting the confused-image of *pepper.bmp*: (a) the decrypted image, which is produced at $b = 3.000000000000001$; (b) histogram of R; (c) histogram of G; (d) histogram of B.

For example, Fig.9 illustrates the sensitivity of the proposed permutation scheme (PPS3DCS) with the secret key b , where as the permuted-image which is shown in Fig.5(a) decrypted using $b = 3.000000000000001$, and the remains secret keys as the same as in the normal case. As can be seen that, even the secret key b is changed a little (10^{-15}), the decrypted image is absolutely different from the original image (*pepper.bmp*).

Therefore anyone can conclude that the proposed permutation scheme (PPS3DCS) is very sensitive to all members of the secret keys, and it can also resist the various attacks based on sensibility.

4.2.2 Information Entropy Analysis

Information entropy [10,16,17] is a common criterion that shows the randomness of the data. Also, entropy and information theory introduced by Robert M. Gray at 2009. two of the most famous formulas of the information entropy are illustrated in formula 9.

$$H(x) = - \sum_{i=0}^{N-1} P(x_i) \text{Lb}(P(x_i)) \quad (9)$$

That N is the number of gray level in the color's channel of the image, x is the total number of symbols, $x_i \in x$, where $P(x_i)$ represents the probability of occurrence of x_i , and Lb denotes the base 2 logarithm.

For an ideal random image, the value of information entropy is 8. The predictability of the method decreases when the information entropy tends to the ideal value (8) [16].

Table 4: Results of Information Entropy analysis for the permuted images of *flower.bmp* and *pepper.bmp* by applying the PPS3DCS .

The Information Entropy $H(x)$			
	R	G	B
<i>Flower.bmp</i>	7.7531	7.9175	7.6624
<i>pepper.bmp</i>	7.6704	7.4326	7.4170

From Table 4 and Fig. 10, all the results of information entropy $H(x)$ for both the images, which are permuted (confused) by PPS3DCS are very close to the ideal value. So these results mean that the permuted-images are close to a random source and the proposed permutation scheme (PPS3DCS) is secure against entropy attack.

Also, from Table 4 and Fig. 10, the information entropy analysis $H(x)$ illustrates the results for the low frequencies colors image (*flower.bmp*) better than the results for the other image.

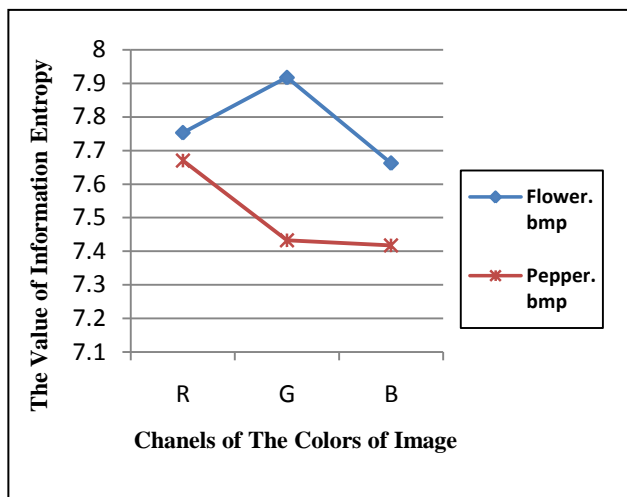


Fig. 10 Values of information entropy analysis for permuted images of *flower.bmp* and *pepper.bmp*

5. Conclusion

In this paper, a new permutation scheme (PPS3DCS) is proposed for colored-images encryption based on Chen's chaotic system. PPS3DCS is the permutation algorithm for shuffling the locations of pixels of the images. The proposed permutation scheme (PPS3DCS) is applied on two different colored-image. The experimental results and analysis show that the proposed scheme (PPS3DCS) is very good algorithm and has high security, where as the proposed permutation scheme (PPS3DCS) has merits: 1) its results with all tests of statistical analysis are excellent. 2) it is very sensitive to all members of the secret keys. 3) its results of information entropy analysis tests are excellent, because these are very closed to the ideal value. As demonstrated in the simulation and its results, the proposed permutation scheme (PPS3DCS) has high encryption quality, and it is suitable to provides an efficient and secure way for the colored-image encryption.

References

- [1] Di X, Xiaofeng L, Pengcheng W. Analysis and improvement of a chaos-based image encryption algorithm. *Chaos, Solitons and Fractals*, Vol.40, No.5, 2009, pp. 2191-2199.
- [2] Xin Ma, Chong Fu, Wei-min Lei, Shuo Li. A novel chaos-based image encryption scheme with an improved permutation process. *IJACT*, Vol.3, No.5, 2011, pp.223-233.
- [3] Dongming Chen, Yunpeng Chang. A novel image encryption algorithm based on Logistic maps. *AISS*, Vol. 3, No.7, 2011, pp.364-372.
- [4] Zhang LH, Liao XF, Wang XB. An image encryption approach based on chaotic maps. *Chaos, Solitons & Fractals*, Vol. 24, 2005, pp. 759-765.
- [5] Wong KW. A fast chaotic cryptography scheme with dynamic look-up table. *Phys Lett A*, Vol. 298, 2002, pp. 238-242.

- [6] Pareek NK, Patidar V, Sud KK. Discrete chaotic cryptography using external key. *Phys Lett A*, Vol. 309, 2003, pp.75-82.
- [7] Guan ZH, Huang FJ, Guan WJ. Chaos-based image encryption algorithm. *Phys Lett A*, Vol. 346, 2005, pp.153-157.
- [8] Lian SG, Sun J, Wang Z. A block cipher based on a suitable use of chaotic standard map. *Chaos, Solitons and Fractals*, Vol. 26, No. 1, 2005, pp.117-129.
- [9] Feng Y, Li LJ, Huang F. A symmetric image encryption approach based on line Maps. In: *Proc ISSCAA2006*, 2006, pp. 1362-67.
- [10] Huibin Lu, Xia Xiao. A Novel Color Image Encryption Algorithm Based on Chaotic Maps. *Advances in information Sciences and Service Sciences (AISS)*, Vol. 3, No. 11, 2011.
- [11] Guanrong Chen, Yaobin Mao, Charles K. Chui. A symmetric image encryption scheme based on 3D chaotic cat maps. *Chaos, Solitons & Fractals*, Vol. 21, 2004, pp. 749–761.
- [12] Xuedi Wang, Lixin Tian, Liqin Yu. Linear Feedback Controlling and Synchronization of the Chen's Chaotic System. *International Journal of Nonlinear Science*, Vol.2, No.1, 2006, pp. 43-49.
- [13] Cahit Cokal, Ercan Solak. Cryptanalysis of a chaos-based image encryption algorithm. *Elsevier Physics Letters A*, Vol. 373, 2009, pp. 1357–1360.
- [14] Tianshou Zhou, Yun Tang, And Guanrong Chen. Chen's Attractor Exists. *International Journal of Bifurcation and Chaos*, Vol. 14, No. 9, 2004, pp. 3167-3177.
- [15] Osama M. Abu Zaid, Nawal A El-fishawy, E M Nigm and Osama S Faragallah. A Proposed Encryption Scheme based on Henon Chaotic System (PESH) for Image Security. *International Journal of Computer Applications*, USA, Vol. 61, No. 5, 2013, pp. 29-39.
- [16] M. Sabery.K, M. Yaghoobi. A New Approach for Image Encryption Using Chaotic Logistic Map. *IEEE Computer Society, ICACTE*, 2008, pp. 585-590.
- [17] Zhiliang Zhu, Wei Zhang, Kwok-wo Wong, Hai Yu. A chaos-based symmetric image encryption scheme using a bit-level permutation. *Information Sciences*, Vol.181, No.6, 2011, pp.1171-1186.

E. M. Nigm received the Ph.D. degree in Mathematical Statistics, Mathematics Dept., the faculty of sciences, Zagazig university, zagazig, Egypt, in 1990. Now he is professor in the mathematics Dept., Faculty of sciences, Zagazig university. His research interest includes Mathematical Statistics and computer sciences. Now he directed her research interests to the developments of security algorithms. he has served as a reviewer for many national and international journals and conferences.

Osama M. Abu Zaid received B.Sc. from the faculty of science, Menoufia University, Egypt in 2000. He is working as a network manager in Menoufia University. He received the M.Sc. degree in data security from Faculty of sciences, Menoufia university, Egypt, in 2005. Now he is lecturer in Faculty of computer sciences and information, Al-Jouf university, KSA. He is working for his Ph.D. He is interested in multimedia security over wired and wireless networks, and he registered the Ph.D. in Faculty of sciences, Zagazig university, Egypt .

Nawal A. El-Fishawy received the Ph.D. degree in mobile communications the faculty of Electronic Eng., Menoufia university, Menouf, Egypt, in collaboration with Southampton university in 1991. Now she is the head of Computer Science and Engineering Dept. , Faculty of Electronic Eng. Her research interest includes computer communication networks. Now she directed her research interests to the developments of security over wireless communications networks, and encryption algorithms. She has served as a reviewer for many national and international journals and conferences. Also she participated in many technical program committees of major international conferences in wireless communications.

CRYPTOSYSTEM ALGORITHM BASED ON CHAOTIC SYSTEMS FOR ENCRYPTING COLORED IMAGES

Osama M. Abu Zaid¹, Nawal A. El-Fishawy², and E. M. Nigm³

¹Ph. D. Researcher in Computer Science, Department of Mathematics, Faculty of Sciences, Zagazig University, Egypt.

²Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Egypt.

³Department of Mathematics, Faculty of Sciences, Zagazig University, Egypt.

Abstract

In this paper, a cryptosystem algorithm based on Chen's and Henon chaotic systems will be introduced and discussed. A proposed cryptosystem algorithm will be designated as CACHS. CACHS meet the requirements of secure image transfer. It will be applied on medium color's frequencies of colored-image. CACHS contains confusion and diffusion algorithms. Confusion algorithm based on Chen's is used to shuffle the positions of pixels of the colored plain-image. Diffusion algorithm based on mixing of Henon and Chen's chaotic systems is used to change the values of pixels of the shuffled-image. CACHS will be applied on the three colors channels of the colored image with two modes of operations ECB and CBC. The results of several experimental, statistical analysis, key sensitivity tests, NPCR and UACI analysis, and information entropy analysis will show that CACHS is a well algorithm to provides an efficient and secure approach for encrypting the colored images.

Keywords Image encryption; Chen's chaotic map; Henon chaotic system; and Modes of operations.

1. Introduction

In recent years, with information and communication technologies developing, how to protect the delivered information transmission over the networks from attacking has become a vital issue [1]. This age of communications revolution which necessitates multimedia transmission in a secure manner. encryption is important in transferring image through the communication networks to protect it against reading, alteration of its content, adding false information, or deleting part of its content.

Chaotic maps are very complicated nonlinear dynamic systems, which are applied in the field of figure correspondence and encryption [2-4], because they are very sensitive to initial conditions and can generate good pseudorandom sequences.

Chaotic systems have many important properties, such as the sensitive dependence on initial conditions and system parameters, pseudorandom property, non-

periodicity and topological transitivity, etc. Most properties meet some requirements such as diffusion and mixing in the sense of cryptography[5].

Traditional cryptosystem algorithms such as Data Encryption Standard (DES) are designed with good confusion and diffusion properties [6]. These two properties can also be found in chaotic systems which are usually ergodic and are sensitive to system parameters and initial conditions.

Traditional image encryption algorithm such as data encryption standard (DES), has the weakness of low-level efficiency when the image is large [7,8]. The chaos-based encryption has suggested a new and efficient way to deal with the intractable problem of fast and highly secure image encryption [9].

Recently, a number of chaos-based encryption schemes have been proposed. Some of them are based on one-dimensional chaotic maps and are applied to data sequence or document encryption [10,11]. For image encryption, two-dimensional (2D) or higher-dimensional chaotic maps are naturally employed as the image can be considered as a 2D array of pixels [12-14]. The colored image consist of three 2D arrays of pixels for the color channels R, G, and B.

This paper will introduce a proposed encryption algorithm for colored images based on intermixture of 2D chaotic map system (Hennon chaotic system) and 3D chaotic map system (Chen's chaotic system). A proposed encryption algorithm will be designated in this paper as (CACHS). An encryption process of the proposed algorithm (CACHS) contains permutation and substitution procedures; so it has the benefits of both of them. Permutation procedure based on Chen 's chaotic system is used to shuffle the positions of pixels of the colored plain-image. Substitution procedure based on mixing of Chen's chaotic system and Hennon chaotic system is used to change the values of pixels of the shuffled-image. CACHS will be applied on Red, Green,

and Blue channels of the colored-image with two modes of operations; Electronic Code Book (ECB) mode, and Cipher Block Chaining (CBC).

This paper is organized as follows. Section 2 presents a brief overview on the chaotic systems which are used in the work. After this, Section 3 discuss a proposed algorithm (CACHS). Section 4 discuss the experimental results and analysis. In the final, Section 5 presents the final conclusion of the paper.

2. The Chaotic Systems

In this section, a brief overview on the two chaotic systems which are used in this work is introduced. These chaotic maps systems are Henon chaotic system and Chen's chaotic system.

2.1 Henon Chaotic System

The Henon map is a discrete-time dynamical system. It is one of the most studied examples of dynamical systems that exhibit chaotic behavior. The Henon map takes a point (x_i, y_i) in the plane and maps it to a new point.

The well-studied Henon map presents a simple two dimensional map with quadratic nonlinearity. This map gave a first example of the strange attractor with a fractal structure. Because of its simplicity, the Henon map easily lends itself to numerical studies. Thus a large amount of computer investigations followed. Nevertheless, the complete picture of all possible bifurcations under the change of the parameters a and b is far from completion. Where $a = 0.3, b \in [1.07, 1.4]$. If one chooses $a=0.3, b=1.4$, the system is chaotic, subsequently This feature is very useful in image encryption. [15,16]

Formula 1 illustrates the two equations of Henon chaotic map system.

$$\begin{aligned} x_{i+1} &= 1 - ax_i^2 + y_i \\ y_{i+1} &= bx_i \end{aligned} \quad (1)$$

2.2 Chen's Chaotic System

Chen's chaotic map system is described by formula 2 which illustrates a set of the three differential equations of Chen's chaotic map system. [1,7,17,18]

$$\begin{cases} x = a(y_0 - x_0) \\ y = (c - a)x_0 - x_0z_0 + cy_0 \\ z = x_0y_0 - bz_0 \end{cases} \quad (2)$$

where $a > 0, b > 0$ and c such that $(2c > a)$ are parameters of the system [19]. Chen's system is chaotic when the parameters have the values; $a=35, b=3$ and $c \in [20, 28.4]$.

When $a=35, b=3$, and $c=28$. It has been experienced that Chen's chaotic system is relatively difficult to control as compared to the Lorenz system due to the prominent three-dimensional and complex dynamic property[1]. Recently, the study about Chen's chaotic map system has attracted many researchers' attention.

In [1], authors obtained on very good performance for Chen's chaotic map at the parameters $a=35, b=3, c=28$, the initial values $x_0 = 0, y_0 = 1, z_0 = 0$, and $h = 0.055555$ such that h is the step of the sequence.

3. A Proposed Encryption Algorithm (CACHS)

In this section, the proposed encryption algorithm based on Chen's and Henon chaotic systems (CACHS) is presented. The proposed algorithm (CACHS) consists of the encryption scheme and the decryption scheme. In this part of the paper the encryption scheme only is discussed because The decryption scheme is the reverse technique of the encryption scheme.

To resist statistical analysis, Shannon suggests that confusion and diffusion should be utilized in any cryptosystem [1]. The encryption scheme of the proposed algorithm (CACHS) consists of two algorithms, the first is permutation and confusion algorithm, and the second is substitution and diffusion algorithm.

3.1 Permutation and Confusion Algorithm

Permutation and confusion algorithm is the first part of designing of the encryption scheme of the proposed algorithm (CACHS). It is designed to permute the positions of the pixels of the image, i.e. shuffling the positions of pixels of the image. This algorithm consists of five steps of operations as following:

Step1: Obtain the $a1, a2$ and $a3$ matrixes (the three color components Red, Green and Blue) of the color image of size $m \times n \times 3$, respectively. $a1$ represents $m \times n$ matrix for the red, $a2$ represents $m \times n$ matrix for the green, and $a3$ represents $m \times n$ matrix for the blue. Afterwards, each color's matrix (including $a1, a2$ and $a3$) is reshaped by Matlab into one dimension matrix (vector) of integers within $\{0, 1 \dots 255\}$, wherein length of the vector is $si = m \times n$. Then, the so obtained three vectors ($aa1, aa2$, and $aa3$) represent the plaintext which will be permuted and encrypted.

Step2: Obtain the sequences XX , YY , and ZZ (1-D matrixes) as in formula 3 which are generated by Chen's chaotic system at $a = 35$, $b = 3$, $c = 28$, the initial values $x_0 = 0+k$, $y_0 = 1+k$, $z_0 = 0+k$, and $h = 0.055555$.

$$\begin{aligned} XX(i) &= \text{floor}(x) \text{ MOD } 256; \\ YY(i) &= \text{floor}(y) \text{ MOD } 256; \\ ZZ(i) &= \text{floor}(z) \text{ MOD } 256; \end{aligned} \quad (3)$$

Where i is from 1 to si . Values of x , y , and z are obtained from the three equations of Chen's chaotic system in formula 2. k is obtained by formula 4, where it is used to modify the keys in the proposed algorithm.

$$k = (k1+k2+k3)/10^13 \quad (4)$$

Formula 5 is employed to generate the values of $k1$, $k2$ and $k3$ which are used to obtain k .

$$\begin{aligned} k1 &= \sum_{i=1}^m \sum_{j=1}^n \mathbf{a1}(i, j) \\ k2 &= \sum_{i=1}^m \sum_{j=1}^n \mathbf{a2}(i, j) \\ k3 &= \sum_{i=1}^m \sum_{j=1}^n \mathbf{a3}(i, j) \end{aligned} \quad (5)$$

Step3: The matrixes XX , YY , and ZZ are sorted in descending sort by using Matlab function (sort). The Matrixes $XX1$, $YY1$, and $ZZ1$ are produced from sorting of the matrixes XX , YY , and ZZ respectively.

For example, let suppose $XX=[1 \ 3 \ 4 \ 10 \ 9 \ 5 \ 2 \ 8 \ 155 \ 255]$, after apply the function of descending sort; the result is $XX1=[255 \ 155 \ 10 \ 9 \ 8 \ 5 \ 4 \ 3 \ 2 \ 1]$. In position expression; the positions $[1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10]$ shifted to the positions $[10 \ 8 \ 7 \ 3 \ 4 \ 6 \ 9 \ 5 \ 2 \ 1]$.

Step4: The reshaped matrixes $aa1$, $aa2$ and $aa3$ are rearranged respectively according to the position of XX in $XX1$, the position of YY in $YY1$, and the position of ZZ in $ZZ1$. The sequences ar , ag , and ab which are obtained from rearranging process of $aa1$, $aa2$, and $aa3$ respectively.

For example, let suppose $aa1=[125 \ 56 \ 90 \ 42 \ 50 \ 220 \ 120 \ 255 \ 65 \ 35]$, according to the position of XX in $XX1$ as in example of step3; the result is $ar = [35 \ 65 \ 42 \ 50 \ 255 \ 220 \ 90 \ 56 \ 120 \ 125]$.

Step5: Obtain the matrixes arp , agp , and abp (the permuted matrixes of the color's matrixes $a1$, $a2$, and $a3$), which are produced respectively by reshaping the sequences ar , ag , and ab from one dimension to the matrixes of two dimension $m \times n$.

According to the confusion algorithm, the position of any pixel in $a1$, $a2$, or $a3$ is different with its position in arp , agp , or abp respectively, which will lead to be strong for the attacks.

3.2 Substitution and Diffusion Algorithm

Substitution and diffusion algorithm is the second part of designing of the encryption scheme of the proposed algorithm (CACHS). It is designed to encrypt the pixels of the image, i.e. changing values of the pixels of the image. Here this algorithm is applied on the permuted image which is produced from the permutation and confusion algorithm in the previous section. This algorithm consists of seven steps of operations as following:

Step1: There are three sequences XX , YY and ZZ of size $m \times n$ which are generated by Chen's chaotic system and are used to permute the $a1$, $a2$ and $a3$ matrixes of the plain-image. Also, there are arp , agp , and abp matrixes of colors of the permuted image which is produced from the permutation procedure according to steps of the previous section.

Step2: The Henon chaotic system is converted into one dimensional chaotic system [30]. The one dimensional Henon chaotic system is defined as in formula 6:

$$w_{i+2} = 1 - aw_{i+1}^2 + bw_i \quad (6)$$

Obtain w_2 , where the initial value $w_0 = 0.01$, and the initial value $w_1 = 0.02$. values of parameters a , and b are the same values of a , and b for Chen's chaotic system.

Step3: The Chen's chaotic system is defined as in the following formula:

$$\begin{aligned} x_2 &= a(y_1 - x_1) \\ y_2 &= (c - a)x_1 - x_1z_1 + cy_1 \\ z_2 &= x_1y_1 - bz_1 \end{aligned} \quad (7)$$

Obtain x_2 , y_2 , and z_2 , where values of the parameters are $a = 35$, $b = 3$, $c = 28$. Also, The three initial values are $x_1 = XX(100)$, $y_1 = YY(500)$ and $z_1 = ZZ(800)$ which are generated by the Chen's chaotic system.

Step4: Obtain two sequences (1-D matrix) ARH and ARC of size $si = m \times n$, where ARH is generated by Henon chaotic system according to the equations in formula 8, and ARC is generated by Chen's chaotic system according to the equations in formula 9. Where i is the variable of the counter for loop, i.e. $i = 1, \dots, si$ at value of the step of the counter is three. And in formula 9 the constant is adopted equal to 10^{14} .

$$\begin{aligned} ARH(i) &= \text{floor}(w_0 * z_1) \text{ MOD } 256; \\ ARH(i+1) &= \text{floor}(w_1 * x_1) \text{ MOD } 256; \\ ARH(i+2) &= \text{floor}(w_2 * y_1) \text{ MOD } 256; \end{aligned} \quad (8)$$

$$\begin{aligned}
 ARC(i) &= \text{floor}((\text{abs}(x_2) * ARH(i) - \text{floor}(\text{abs}(x_2))) * \\
 &\quad \text{constant} \text{ MOD } 256); \\
 ARC(i+1) &= \text{floor}((\text{abs}(y_2) * ARH(i+1) - \text{floor}(\text{abs}(y_2))) * \\
 &\quad \text{constant} \text{ MOD } 256); \\
 ARC(i+2) &= \text{floor}((\text{abs}(z_2) * ARH(i+2) - \text{floor}(\text{abs}(z_2))) * \\
 &\quad \text{constant} \text{ MOD } 256);
 \end{aligned} \tag{9}$$

At the end of each loop of the counter, the initial values w_0 , w_1 , x_1 , y_1 , and z_1 are changed according to the following formula:

$$\begin{aligned}
 x_1 &= x_2 * w_0; \\
 y_1 &= y_2 * w_1; \\
 z_1 &= z_2 * w_2; \\
 w_0 &= w_1; \\
 w_1 &= w_2;
 \end{aligned} \tag{10}$$

Step5: Obtain three sequences (1-D matrixes) XXC , YYC , and ZZC of size $si = m \times n$, where these sequences based on the sequence ARC which is produced in step4 and the values of $k1$, $k2$, and $k3$ which are produced in step2 of the previous subsection. XXC , YYC , and ZZC are generated according to the equations in formula 11.

$$\begin{aligned}
 XXC(i) &= ((k1+k2)*ARC(i)) \text{ MOD } 256; \\
 YYC(i) &= ((k2+k3)*ARC(i)) \text{ MOD } 256; \\
 ZZC(i) &= ((k3+k1)*ARC(i)) \text{ MOD } 256;
 \end{aligned} \tag{11}$$

Step6: XXC , YYC and ZZC are changed based on exclusive OR operation for themselves with the sequence ARH which is produced in step4. A new sequences XXC , YYC , and ZZC are generated according to the equations in formula 12.

$$\begin{aligned}
 XXC(i) &= XXC(i) \oplus ARH(i); \\
 YYC(i) &= YYC(i) \oplus ARH(i); \\
 ZZC(i) &= ZZC(i) \oplus ARH(i);
 \end{aligned} \tag{12}$$

Step7: Then the matrixes of colors of the encrypted image can be obtained by the following formula:

$$\begin{aligned}
 EIMR(i, j) &= arp(i, j) \oplus XXC(t); \\
 EIMG(i, j) &= agp(i, j) \oplus YYC(t); \\
 EIMB(i, j) &= abp(i, j) \oplus ZZC(t);
 \end{aligned} \tag{13}$$

Where arp , agp , and abp are the color's matrixes of the permuted-image which are generated in step5 of the previous subsection. Also, i is the first dimension of the matrixes where $i = 1, \dots, m$ and j is the second dimension of the matrixes where $j = 1, \dots, n$. Also, $t = 1, \dots, si$, where $si = m \times n$.

According to all previous steps, it is clear that the generation of the key streams depends on the plaintext through all the color components, and every pixel value of the encryption matrix EIM includes the information of all the color components, i.e. the diffusion has been taken full advantages. These features strengthen the proposed encryption algorithm security.

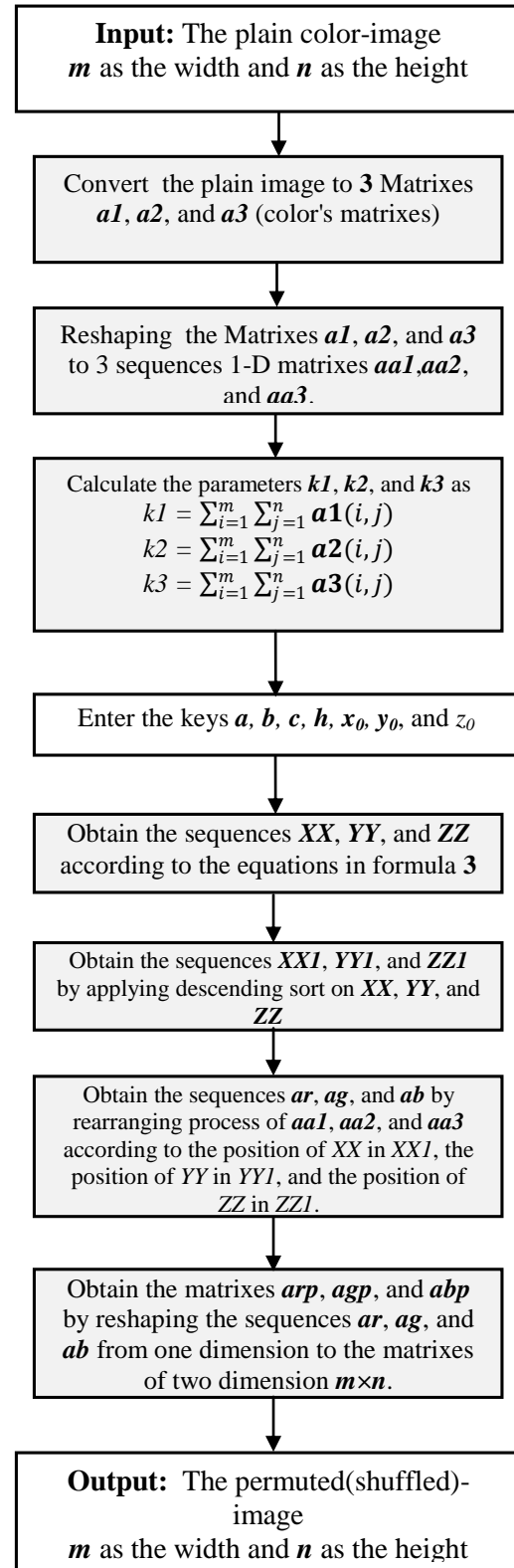


Fig. 1 The Data-Flow diagram for a permutation (Confusion) algorithm of encryption scheme of CACHS

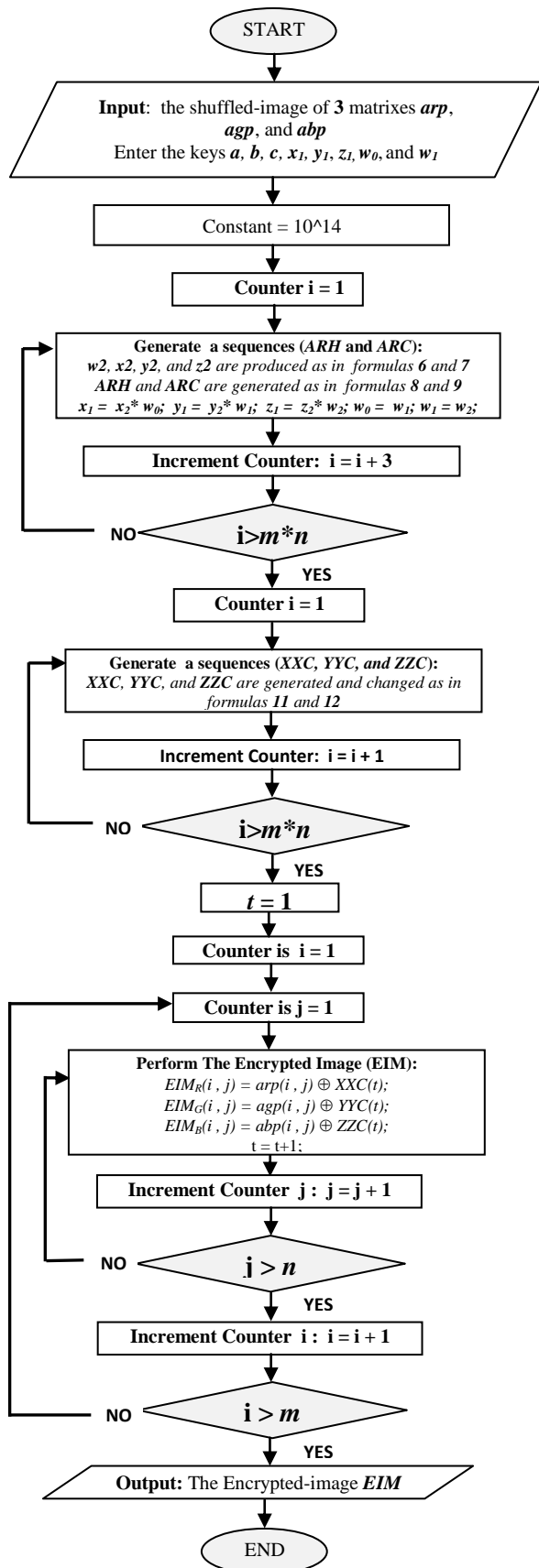


Fig. 2 The Flow-Chart for a substitution (Diffusion) algorithm of encryption scheme of CACHS

4. Experimental Results and Analysis

In this paper, a practical programs of a proposed encryption algorithm (CACHS) with the modes of operations and a practical programs of all experimental and security analysis tests are designed by MATLAB 7.0 on windows 7 system on Laptop computer with Intel CORE I₃ Processor, 3.0 GB RAM, and 320 GB Hard Disk. All programs have been applied on the colored-image which have high frequency of colors (*fruit.bmp*) as a plain-image of the size 120×120 pixels, which are shown in Fig. 3(a).

4.1 Statistical Analysis

To examine the quality of encryption and the stability via statistical attacks, the histogram is calculated for all color's channels *R*, *G*, *B* of the plain-images, correlation coefficient (CC) between each of color's channels *R*, *G*, *B* of the plain-image and the corresponding channels of the encrypted-image, the correlation analysis of two adjacent pixels with the directions horizontal (HC), vertical (VC), and diagonal (DC) for all color's channels *R*, *G*, *B* of the encrypted-image.

4.1.1 Histogram Analysis

The plain colored-image (*fruit.bmp*) of the size 120×120 pixels are shown in Fig.3(a), and the histogram for *R*, *G*, *B* of this image is shown in Fig.3(b, c, d).

The application of the proposed algorithm (CACHS) on this image has two sequent steps; first is permutation (confusion) algorithm and second is substitution (diffusion) algorithm.

Figure 4(a) show the shuffled-image for *fruit.bmp* which are produced from applying the permutation (confusion) algorithm. The histogram for *R*, *G*, *B* of this image is shown in Fig.4(b, c, d) which are the same histograms of the plain-image.

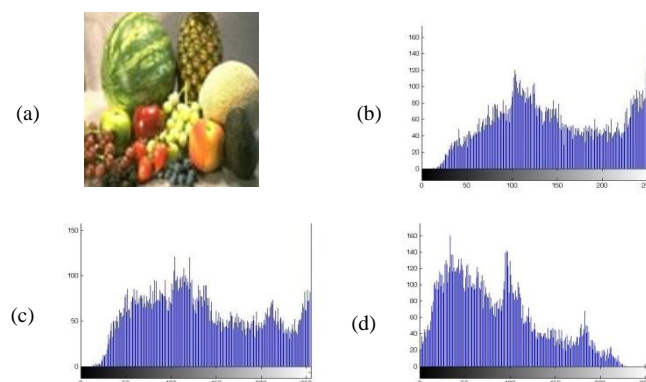


Fig. 3 The plain colored-image and its histogram: (a) the image (*fruit.bmp*); (b) histogram of *R*; (c) histogram of *G*; (d) histogram of *B*.

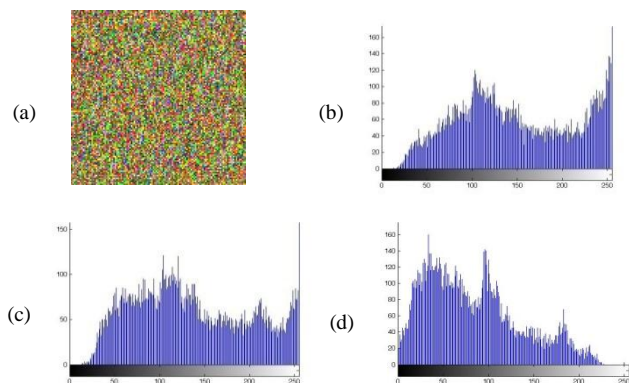


Fig. 4 The shuffled-image for fruit.bmp and its histogram: (a) the shuffled-image; (b) histogram of R; (c) histogram of G; (d) histogram of B.

Figure 5(a) illustrates the encrypted-image for *fruit.bmp* which is produced from applying the proposed cryptosystem algorithm (CACHS) with ECB mode. The histogram for R, G, B of the encrypted-image is shown in Fig. 5(b, c, d).

Figure 6(a) illustrates the encrypted-image for *fruit.bmp* which is produced from applying the proposed cryptosystem algorithm (CACHS) with CBC mode. The histogram for R, G, B of the encrypted-image is shown in Fig. 6(b, c, d).

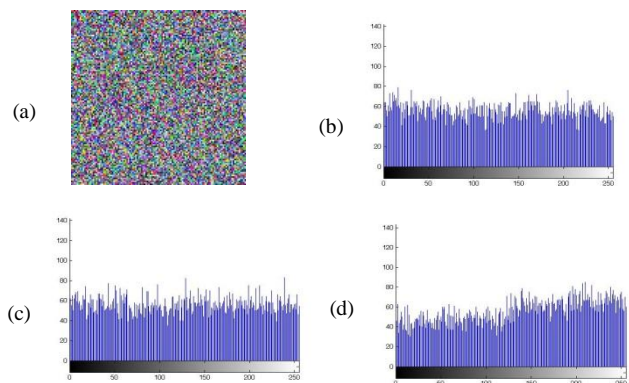


Fig. 5 The encrypted-image for fruit.bmp which are produced by applying CACHS with ECB mode: (a)the encrypted-image; (b) histogram of R; (c) histogram of G; (d) histogram of B.

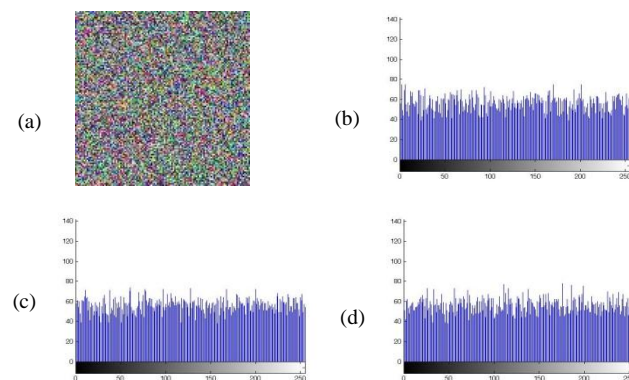


Fig. 6 The encrypted-image for fruit.bmp which are produced by applying CACHS with CBC mode: (a)the encrypted-image; (b) histogram of R; (c) histogram of G; (d) histogram of B.

From all previous figures of histograms, as anyone can see, the histogram of the encrypted-image is fairly uniform and is significantly different from that of the plain-image. The proposed algorithm (CACHS) is a complicated and very good procedure for disguise any countenance of the image. Also, anyone can observe, the proposed algorithm (CACHS) is qualification for encrypting the colored-images.

4.1.2 Correlation Coefficient Analysis

The correlation coefficient equals one if they are highly dependent, i.e. the encryption process failed in hiding the details of the plain-image. If the correlation coefficient equals zero, then the plain-image and its encryption are totally different. So, success of the encryption process means smaller values of the CC [20]. The CC is measured by formula 14:

$$CC = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^N (x_i - E(x))(y_i - E(y))}{\sqrt{\sum_{i=1}^N (x_i - E(x))^2} \sqrt{\sum_{i=1}^N (y_i - E(y))^2}} \quad (14)$$

$$\text{where } E(x) = \frac{1}{N} \sum_{i=1}^N x_i$$

where x and y are gray-scale pixel values of the plain and encrypted images. The CC is measured for each color's channel (R, G, B) of any colored-image.

Tables 1, illustrates that the proposed cryptosystem algorithm (CACHS) achieves very small values (near to zero) of CC with all modes of operations for the colored-image, so a CACHS is a complicated and a good algorithm for encrypting the images. Also, the results of CC for CBC mode is better than the results for ECB mode.

Table 1: Results of CC analysis for encrypting *fruit.bmp* by CACHS with the modes.

Modes	CC for encrypting <i>fruit.bmp</i>		
	R	G	B
ECB	0.0012	-0.0029	-0.0069
CBC	0.0014	-0.0020	0.00095

4.1.3 Correlation Analysis of Two Adjacent Pixels

It is well known that the adjacent pixels of an image have very high correlation coefficients in horizontal, vertical and diagonal directions. The following formulas is employed to test the correlation between two horizontally adjacent pixels (designed as **HC**), two vertically adjacent pixels (designed as **VC**), and two diagonally adjacent pixels (designed as **DC**), respectively, in plain images and encrypted images, the following procedure was carried out. First, select 900 pairs of two adjacent pixels from an image. Then,

calculate the correlation coefficient r_{xy} of each pair by using the following formulas [1,7]:

$$E(x) = \frac{1}{N} \sum_{i=1}^N x_i, D(x) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2 \quad (15)$$

$$cov(x, y) = E(x - E(x))(y - E(y)) \quad (16)$$

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)} \sqrt{D(y)}} \quad (17)$$

Where x and y denote two adjacent pixels, and N is the total number of duplets (x, y) obtained from the image. Table 2 illustrates the results of HC, VC, and DC analysis for the plain colored-image (*fruit.bmp*).

Table 3 illustrates the results of HC, VC, and DC analysis for the two encrypted-images, which are produced by applying the proposed cryptosystem algorithm (CACHS) on the plain-image with the two modes ECB and CBC.

Table 2: Results of HC, VC, and DC analysis for the plain colored-image *fruit.bmp*.

	The plain image (<i>fruit.bmp</i>)		
	R	G	B
HC	0.9367	0.9433	0.9287
VC	0.9827	0.9812	0.9719
DC	0.9153	0.9010	0.9000

Table 3: Results of HC, VC, and DC analysis for the encrypted images of *fruit.bmp* by applying the CACHS with the modes.

		The encrypted image of (<i>fruit.bmp</i>)		
		R	G	B
HC	ECB	0.0026	0.0272	0.00059
	CBC	-0.0066	0.0062	-0.0751
VC	ECB	-0.00007	-0.0031	-0.0068
	CBC	-0.0396	-0.00042	0.0039
DC	ECB	-0.0024	0.0089	-0.0185
	CBC	-0.0614	0.0069	-0.0371

According to Table 2, anyone can observe, the results of HC, VC, and DC for the correlation analysis of two adjacent pixels for the plain colored-image are approach to 1, implying that high correlation exists among pixels.

According to Table 3, the results of HC, VC, and DC for the correlation analysis of two adjacent pixels for the encrypted-image with both of two modes are approach to 0, implying that no detectable correlation exists among pixels. Therefore the proposed cryptosystem algorithm (CACHS) can protect the encrypted-image from statistical attacks. Also, from Table 3, the results sometimes better with ECB than CBC and other sometimes the converse is actualize.

4.2 Security Analysis

A good encryption algorithm should resist most kinds of known attacks, also it must be achieves sensitive to any little change in the plain-text or secret keys, large enough in the key space to make brute-force attacks infeasible, and a good values for the information entropy analysis.

In the proposed cryptosystem algorithm (CACHS), the parameters $a, b, c,$ and $h,$ the initial values $x_0, y_0, z_0, w_0, w_1, x_1, y_1,$ and z_1 are used as a secret keys. The key space is large enough to resist all kinds of brute-force attacks.

4.2.1 The Plain-text Sensitivity Analysis

A very vital relationship between the plain-image and the encrypted-image may be revealed [1]. If a significant change in the encrypted-image can be caused by a trivial change in the plain-image by means of diffusion and confusion, then the algorithm would make differential attacks practically useless. In order to test the influence of a one pixel change on the plain-images encrypted by the proposed algorithm (CACHS), NPCR (Number of Pixels Change Rate) and UACI (Unified Average Changing Intensity) are used. NPCR and UACI are computed by the following formulas [1, 20]:

$$NPCR = \frac{\sum_{i,j} D(i,j)}{M \times N} \times 100\% \quad (18)$$

$$UACI = \frac{1}{M \times N} \sum_{i,j} \frac{|E1(i,j) - E2(i,j)|}{255} \times 100\% \quad (19)$$

$$\text{Where } D(i, j) = \begin{cases} 0, & E1(i, j) = E2(i, j) \\ 1, & E1(i, j) \neq E2(i, j) \end{cases}$$

This test needs two plain-images: the plain-image and the other image obtained by changing one pixel value of the plain-image. the two images are encrypted by a proposed cryptosystem algorithm (CACHS) with the same keys to generate the corresponding encrypted-images $E1$ and $E2$. Where the grey values of the pixel at position (i, j) of $E1$ and $E2$ are denoted as $E1(i, j)$ and $E2(i, j)$ respectively; M and N are width and height of the encrypted-image. $D(i, j)$ is determined by $E1(i, j)$ and $E2(i, j)$.

Table 4: Results of NPCR and UACI analysis for the encrypted images of *pepper.bmp* by applying the CACHS with the modes.

		For The Encrypted images of (<i>fruit.bmp</i>)			
		R	G	B	AVG.
NPCR %	ECB	99.604	99.660	99.611	99.625
	CBC	99.611	99.611	99.604	99.609
UACI %	ECB	33.653	33.778	33.352	33.594
	CBC	33.712	33.143	33.479	33.445

From Table 4, anyone can observe that the results of NPCR and UACI for the two encrypted images with both of modes (ECB, and CBC) are very close to the ideal values ($NPCR=99.609\%$ and $UACI=33.4635\%$)[1], i.e. with the proposed cryptosystem algorithm, A very little change of the plain-image pixel values (one pixel) will lead to a significant change of the encrypted-image.

4.2.2 The Key Sensitivity Analysis

The experimental results demonstrate that the proposed cryptosystem algorithm (CACHS) is very sensitive to the secret keys mismatch. The decrypted image by using CACHS are the same of the original image, where are decrypted by using CACHS with $a=35$, $b=3$, $c=28$, $h=0.055555$, $x_0=0+k$, $y_0=1+k$, $z_0=0+k$, $w_0=0.01$, $w_l=0.02$, $x_l=XX(100)$, $y_l=XX(500)$, and $z_l=XX(800)$ to produce the original image.

The experimental results for applying CACHS on *fruit.bmp* with both of modes demonstrate that the proposed cryptosystem algorithm (CACHS) is very sensitive to the secret keys a mismatch (10^{-14}), b mismatch (10^{-15}), c mismatch (10^{-14}), h mismatch (10^{-16}), x_0 mismatch (10^{-16}), y_0 mismatch (10^{-15}), z_0 mismatch (10^{-14}), w_0 mismatch (10^{-17}), w_l mismatch (10^{-17}), x_l mismatch (10^{-13}), y_l mismatch (10^{-14}), and z_l mismatch (10^{-14}).

For example, Fig.7 illustrates the sensitivity of the proposed algorithm (CACHS) with the secret key w_0 , where as the encrypted-image which is shown in Fig.3(a) decrypted using $w_0=0.01000000000000001$, and the remains secret keys as the same as in the normal case. As can be seen that, even the secret key w_0 is changed a little (10^{-17}), the decrypted image is absolutely different from the original image (*fruit.bmp*).

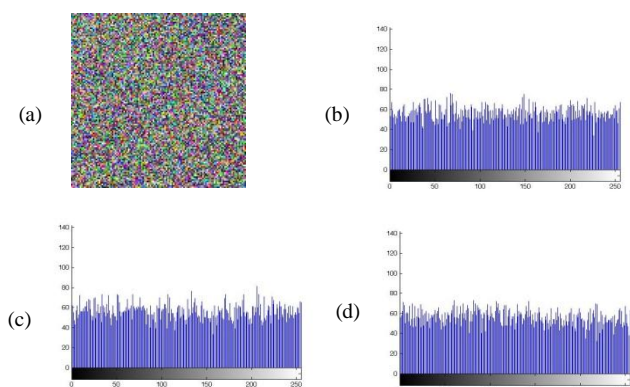


Fig. 7 The sensitivity to the secret key w_0 of CACHS with ECB, for decrypting the encrypted-image of *fruit.bmp*: (a) the decrypted image, which is produced at $w_0=0.01000000000000001$; (b) histogram of R; (c) histogram of G; (d) histogram of B.

Another for example, Fig.8 illustrates the sensitivity of the proposed algorithm (CACHS) with the secret key b , where as the encrypted-image which is shown in Fig.3(a) decrypted using $b=3.000000000000001$, and

the remains secret keys as the same as in the normal case. As can be seen that, even the secret key b is changed a little (10^{-15}), the decrypted image is absolutely different from the original image (*fruit.bmp*).

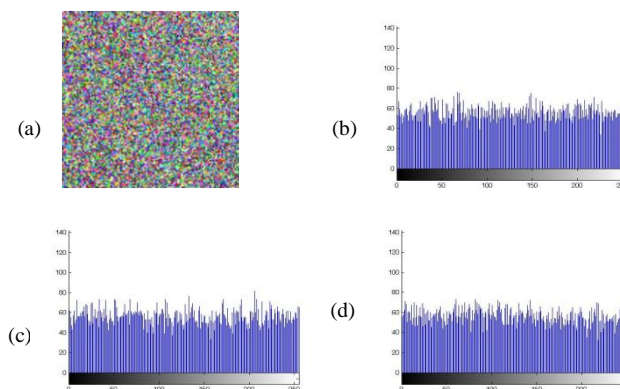


Fig. 8 The sensitivity to the secret key w_0 of CACHS with ECB, for decrypting the encrypted-image of *fruit.bmp*: (a) the decrypted image, which is produced at $b=3.000000000000001$; (b) histogram of R; (c) histogram of G; (d) histogram of B.

Therefore anyone can conclude that CACHS is very sensitive to all members of the secret keys, and it can also resist the various attacks based on sensibility.

4.2.3 Information Entropy Analysis

Information entropy [1,21,22] is a common criterion that shows the randomness of the data. Also, entropy and information theory introduced by Robert M. Gray at 2009. two of the most famous formulas of the information entropy are illustrated in formulas 20 and 21 from [1] and [21] respectively.

$$H(x) = - \sum_{i=0}^{N-1} P(x_i) \text{Lb}(P(x_i)) \quad (20)$$

or

$$H(x) = \sum_{i=0}^{N-1} P(x_i) \text{Log}\left(\frac{1}{P(x_i)}\right) \quad (21)$$

That N is the number of gray level in the color's channel of the image, x is the total number of symbols, $x_i \in x$, where $P(x_i)$ represents the probability of occurrence of x_i , and Lb denotes the base 2 logarithm.

Table 5: Results of Information Entropy analysis for the encrypted image of *fruit.bmp* by applying the CACHS with the modes.

		The Information Entropy $H(x)$		
		R	G	B
<i>fruit.bmp</i>	ECB	7.986	7.984	7.970
	CBC	7.986	7.988	7.987

For an ideal random image, the value of information entropy is 8. The predictability of the method decreases when the information entropy tends to the ideal value (8)

[21]. From Table 5, all the results of information entropy $H(x)$ for the image, which is encrypted by CACHS with both of the modes are very close to the ideal value. So these results mean that the encrypted-image are close to a random source and the proposed cryptosystem algorithm (CACHS) is secure against entropy attack.

Also from Table 5 and Fig. 9, the information entropy analysis $H(x)$ illustrates the results with both of modes ECB and CBC are convergent, but are better with CBC mode than with ECB mode.

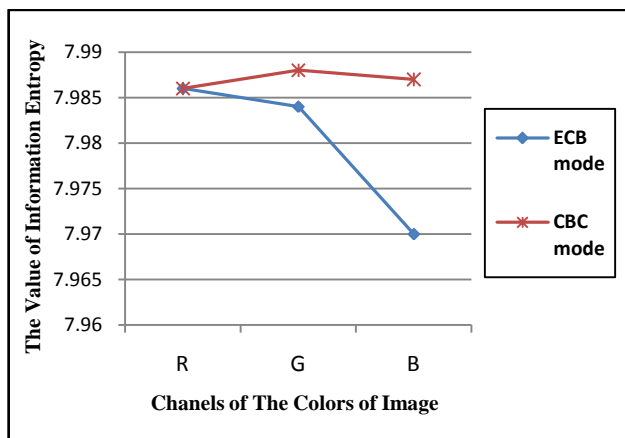


Fig. 9 Values of information entropy analysis for encrypted image of fruit.bmp with the two modes ECB and CBC

5. Conclusion

In this paper, a new cryptosystem algorithm (CACHS) is proposed for colored-images encryption based on Henon chaotic system and Chen's chaotic system. CACHS contains the confusion algorithm for shuffling the locations of pixels of the images, and the diffusion algorithm for encrypting the shuffled-images by changing the values of pixels of the images. The proposed cryptosystem algorithm (CACHS) is applied on the colored-image with two modes of operations ECB and CBC. The experimental results and analysis show that the proposed cryptosystem algorithm (CACHS) is very good encryption Algorithm and has high security, where as the proposed cryptosystem algorithm (CACHS) has merits: 1) its results with all tests of statistical analysis are excellent. 2) it has a large enough key space to resist most kinds of brute force attacks. 3) it is very sensitive to all members of the secret keys. 4) its results of NPCR and UACI tests are excellent, because these are very closed to the ideal values. 5) its results of information entropy analysis tests are excellent, because these are very closed to the ideal value 8. As demonstrated in the simulation and its results, the proposed cryptosystem algorithm (CACHS) has high encryption quality, and it is suitable to provides an efficient and secure way for the colored-image encryption.

References

- [1] Huibin Lu, Xia Xiao, "A Novel Color Image Encryption Algorithm Based on Chaotic Maps," *Advances in information Sciences and Service Sciences (AISS)*, Vol. 3, No. 11, December 2011.
- [2] Di X, Xiaofeng L, Pengcheng W, "Analysis and improvement of a chaos-based image encryption algorithm," *Chaos, Solitons and Fractals*, Vol.40, No.5, 2009, pp. 2191-2199.
- [3] Xin Ma, Chong Fu, Wei-min Lei, Shuo Li, "A novel chaos-based image encryption scheme with an improved permutation process," *IJACT*, Vol.3, No.5, 2011, pp.223-233.
- [4] Dongming Chen, Yunpeng Chang, "A novel image encryption algorithm based on Logistic maps," *AISS*, Vol. 3, No.7, 2011, pp.364-372.
- [5] Zhang LH, Liao XF, Wang XB, "An image encryption approach based on chaotic maps," *Chaos, Solitons & Fractals*, Vol. 24, 2005; pp. 759-765.
- [6] Schneier B. *Cryptography: Theory and Practice*. Boca Raton: CRC Press; 1995.
- [7] Guanrong Chen, Yaobin Mao, Charles K. Chui, "A symmetric image encryption scheme based on 3D chaotic cat maps," *Chaos, Solitons & Fractals* Vol. 21, 2004, pp. 749-761.
- [8] Chiaraluce F, Ciccarelli L, et al. "A new chaotic algorithm for video encryption". *IEEE Trans Consume Electron*, Vol. 48, 2002, pp. 838-843.
- [9] Haojiang Gao, Yisheng Zhang, Shuyun Liang, Dequn Li, "A new chaotic algorithm for image encryption," *Chaos, Solitons and Fractals* , Vol.29 , 2006, pp. 393-399.
- [10] Wong KW, "A fast chaotic cryptography scheme with dynamic look-up table," *Phys Lett A* , Vol. 298,2002, pp. 238-242.
- [11] Pareek NK, Patidar V, Sud KK, "Discrete chaotic cryptography using external key," *Phys Lett A*, Vol. 309,2003, pp.75-82.
- [12] Guan ZH, Huang FJ, Guan WJ, "Chaos-based image encryption algorithm," *Phys Lett A*, Vol. 346,2005, pp.153-157.
- [13] Lian SG, Sun J, Wang Z, "A block cipher based on a suitable use of chaotic standard map," *Chaos, Solitons and Fractals*, Vol. 26, No. 1,2005, pp.117-129.
- [14] Feng Y, Li LJ, Huang F, "A symmetric image encryption approach based on line Maps," In: *Proc ISSCAA2006*, Jan 2006, p. 1362-67.
- [15] R.Raja Kumar, A.Sampath, P.Indumathi, "Enhancement and Analysis of Chaotic Image Encryption Algorithms," *CCSEA 2011, CS & IT 02*, 2011,pp. 143-153.
- [16] M, Sonls, "Once more on Henon map: analysis of bifurcations," *Chaos, Sotilons Fractals*, Vol. 7, No. 12, 1996, pp. 2215-2234.
- [17] Xuedi Wang, Lixin Tian, Liqin Yu, "Linear Feedback Controlling and Synchronization of the Chen's Chaotic System," *International Journal of Nonlinear Science*, Vol.2, No.1, 2006, pp. 43-49.
- [18] Cahit Cokal, Ercan Solak, "Cryptanalysis of a chaos-based image encryption algorithm," *Elsevier, Physics Letters A*, Vol. 373, 2009, pp. 1357-1360.
- [19] Tianshou Zhou, Yun Tang, And Guanrong Chen, "Chen's Attractor Exists," *International Journal of Bifurcation and Chaos*, Vol. 14, No. 9, 2004, pp. 3167-3177.
- [20] Osama Abu M Zaid, Nawal A El-fishawy, E M Nigm and Osama S Faragallah, "A Proposed Encryption Scheme based on Henon Chaotic System (PESH) for Image

Security," International Journal of Computer Applications, USA, Vol. 61, No. 5, 2013, pp. 29-39.

- [21] M. Sabery.K, M. Yaghoobi, "A New Approach for Image Encryption Using Chaotic Logistic Map," IEEE Computer Society, ICACTE, 2008, pp. 585-590.
- [22] Zhiliang Zhu, Wei Zhang, Kwok-wo Wong, Hai Yu, "A chaos-based symmetric image encryption scheme using a bit-level permutation," Information Sciences, Vol.181, No.6, 2011, pp.1171-1186.

Osama M. Abu Zaid received B.Sc. from the faculty of science, Menoufia University, Egypt in 2000. He is working as a network manager in Menoufia University. He received the M.Sc. degree in data security from Faculty of sciences, Menoufia university, Egypt, in 2005. Now he is lecturer in Faculty of computer sciences and information, Al-Jouf university, KSA. He is working for his Ph.D. He is interested in multimedia security over wired and wireless networks, and he registered the Ph.D. in Faculty of sciences, Zagazig university, Egypt .

Nawal A. El-Fishawy received the Ph.D. degree in mobile communications the faculty of Electronic Eng., Menoufia university, Menouf, Egypt, in collaboration with Southampton university in 1991. Now she is the head of Computer Science and Engineering Dept. ., Faculty of Electronic Eng. Her research interest includes computer communication networks. Now she directed her research interests to the developments of security over wireless communications networks, and encryption algorithms. She has served as a reviewer for many national and international journals and conferences. Also she participated in many technical program committees of major international conferences in wireless communications.

E. M. Nigm received the Ph.D. degree in Mathematical Statistics, Mathematics Dept., the faculty of sciences, Zagazig university, zagazig, Egypt, in 1990. Now he is professor in the mathematics Dept., Faculty of sciences, Zagazig university. His research interest includes Mathematical Statistics and computer sciences. Now he directed her research interests to the developments of security algorithms. he has served as a reviewer for many national and international journals and conferences.

New algorithm for automatic visualization of metro map

Somayeh Sobati moghadam

Electrical and computer engineering faculty, Hakim Sabzevari University, Sabzevar, Iran.

Abstract

The metro map is a type of diagram that illustrates transportation network. The automated method for visualisation of graphs with metro map layout is a challenge. In this paper two method for metro map automatic visualization is proposed. These methods use theory of graph and a grid for node coordination. In second method a temperature for graph is considered and the position of node is changed according to simulated annealing. The results show that these algorithms can produce good Metro Maps. The first algorithm has a good result for the graphs with many connections and the second, is better for very dense graphs.

Keywords: Metro map layout, simulated annealing, graph.

1. Introduction

The metro map is a type of diagram that is used for illustrating transportation networks. Designing a schematic map today is still a challenge. Metro map visualizes the interconnections of rail road networks, and people are able to use it quickly. Many people around the world know how to read route maps for metros and buses [1]. The *metro map* can be presented as a graph; it consists of a set of lines which have intersections or overlaps. In all metro maps it is common to consider the graph in which, metro stations are considered as vertices and their interconnections as edges. The lines are straightened and restricted to horizontals, verticals and diagonals at 45°. For forming a good metro map layout, a set of aesthetic criteria should be defined. Traditionally, metro maps are drawn manually [2] and there was no automatic way to produce this diagram; the cartographer must decide where to put the stations and how to draw the lines in the diagram. In complex and large networks it is not simple. In the other hand this kind of diagram could be used as metaphor for abstract entities. The metro map diagram is easy to use and understanding hence it can be use for illustrate the data and relationship between them. For example in [3] it is used for organizing web-based learning resources. There are different types of this usage [4, 5, and 6]. First, we have to know what kind of map is suitable for metro map and what is a “nice” metro map? The map should be as readable and clear as possible without displaying unnecessary details [4]. The lines in the map must be displayed horizontally, vertically or diagonally at 45°. This final layout is called “octilinear” layout. The final map must have as minimum bends as possible without edge crossing and overlapping of labels. Hong and al. [1] proposed five methods. For automatic drawing metro map; they used a modified

embedded spring system. Stott and Rodger [1] are investigating another approach with considering multi criteria optimization. Nollenburg and Wolff [7], is proposing an automated method by the definition of hard and soft constraint.

2. Preliminaries

In this section the main concepts in graph drawing is reviewed. In the state of the art the related work is described and in the final section the new algorithm is proposed.

2.1 Graphs

Graphs are important because they can be used to represent essentially *any* relationship between entities. Graphs visualize the information for the users, and provide important information about the objects. For example, graphs can model a network of roads, with cities as vertices and roads between cities as edges. A nice layout of graph aids user to find immediately the information that he is looking for.

A graph $G = (V, E)$, consists of a set of vertices V , $|V|=n$, and a set of edges E , $|E|=m$. An edge is a pair $e = (u, v)$, $u, v \in V$. Two adjacent vertexes are connected by an edge. The degree of a vertex is the number of edges that incident to this vertices. A path in G is a sequence of distinct vertices of G like (v_1, v_2, \dots, v_t) such that $v_i v_{i+1} \in E$ for $1 < i < t$.

A graph is called planar if it can be drawn in the plan without edge crossing.

In this paper, we need to use another definition of a graph based on lines. So, our problem, a graph $G = (V, E, P)$, will consist of a set of vertex V , a set of edges E , and a set of paths P . In any path there are many nodes with degree two.

2.2 Graph layout

Graph drawing applies topology and geometry to draw two and three-dimensional representation of graphs. Very different layouts can correspond to the same graph. There are different graph layout strategies. Fig.1 shows an example of a graph layout. In *straight line* drawing each edge is drawn as the straight line between the vertices. *Orthogonal layout* the edges are drawn as polygonal chains of horizontal and vertical line segments. We will use the *octilinear* layout for metro map drawing. In octilinear layout all edges are horizontally, vertically or 45°. The advantage of using this layout in comparison of orthogonal layout is that the maximum possible vertex degree increases from 4 to 8. Some drawings are better

than the others. Aesthetic criteria attempt to characterize readability of a layout. Various attempts have been made to specify the readability of a layout that comprises:

- minimize *crossings*
- minimize *area*
- minimize *bends* (in orthogonal drawings)
- minimize *slopes* (in polyline drawings)
- maximize *smallest angle*
- maximize display of *symmetries*.

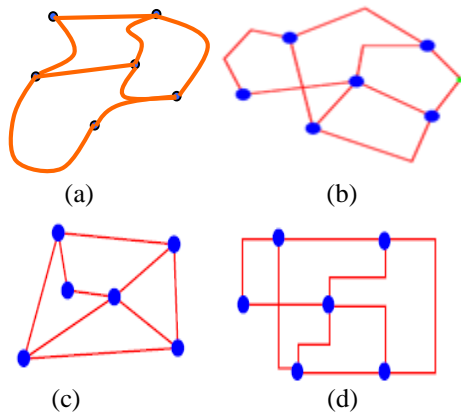


Fig.1 Four layouts of the same graph G. In (a) drawing with simple open curves. In (b) polyline drawing, (c) straight line, (d) orthogonal drawing of G.

In the other hand there are graphical properties for readability of a layout. For example the shape and size and color of vertices; the shape (curve, line), thickness and color of edges. In any algorithmic method of graph drawing these properties are left to the decision of user. Another property that must be taken into account is the graph labeling. The size and place of labels are important to minimize overlaps.

2.3 Simulated annealing

Simulated annealing is a very general optimization method, which stochastically simulates the slow cooling of a physical system. The idea is that there is a cost function H which associates a cost with a proposing change and their accepting or rejecting each change. Having proposed a change we may evaluate the change in H . The proposed change may be accepted or rejected by the criterion; if the cost function decreases the change is accepted. The system soon reaches a state in which none of the proposed changes can decrease the cost function. The simulated annealing is used in our method. In second method, a temperature is calculated for the system, in each iteration, if it is not zero the state of system changes until the temperature reaches at zero.

3. Related works

Automatic drawing a graph with a desired layout is a challenge. For octilinear drawing some methods are proposed. Each method has several advantage and disadvantage.

Nollenberg and Wolff [7] have investigated an algorithm for automatically drawing metro maps. A planar graph G with maximum degree 8 and a set of paths or cycles, L , is considered. A list of soft and hard constraints is specified. Then they have shown that it is NP-hard to decide whether a drawing of G can satisfying all hard constrains. They have presented a mixed- integer linear program (MIP) which finds a drawing that fulfills all hard constraints. This program optimizes a weighted sum of costs corresponding to the soft constraints. They have shown how to include vertex labels in the drawing [7]. The authors have implemented this approach for six real-world examples. Stott and Rodgers in [1] have proposed an automatic method for drawing metro maps. They have considered the metro map as a graph. The graph is embedded on an integer square grid. It is important to mention that more than one node does not share the same grid intersection. In the case of contention for a particular intersection, the node being snapped should be moved to the nearest grid intersection that is vacant [1].

They have implemented a total of eight criteria. Each criterion measures some geometric property of the map such as the length of edges or edges crossings and is weighted. The nodes and labels are repositioned such that the total of the weighted criteria is always reduced. Hill climbing algorithm is based on the total of the weighted criteria, t , which t has given by:

$$t = w_1c_1 + w_2c_2 + w_3c_3 + w_4c_4 + w_5c_5 \quad (1)$$

Which c_1, c_2, c_3, c_4, c_5 denote respectively the number of edge crossings, the edge length, the angular resolution, the line straightness and four-gonality respectively. An initial value of t , is t_0 . The nodes are moved to the location that has the smallest value of t . When selecting a position to move a node there are a number of points that need to be considered: the total value of the weighted criteria, t ; whether or not another node occupies that grid intersection; whether moving the node would occlude other nodes or edges; how far to move the node; whether the distance to move the node is reduced with each iteration (cooling); and whether the cyclic ordering of edges incident to a node would change [1]. In each movement for a node, the initial value of t , t_0 , is calculated. A set of locations, T , is remembered for each movement where $t < t_0$. The node is moved to the location in T that has the smallest value of t_0 . Force directed algorithms are used for graph drawing but they introduce a lot of edge crossing. F. Bertault [8] has proposed an algorithm based on a force-directed approach. A force $F(v)$, for each node v , is computed. Then each node is moved in the direction of $F(v)$. Three kinds of forces are considered: the attraction forces between nodes, the repulsion forces between each pair of nodes, and the repulsion forces between nodes and edges [8]. For each node, v a zone $Z(v)$ is associated that preserves the crossing properties between edges. This zone is divided into eight zones. For each node the size of the zones is updated. PrEd algorithm produces drawings which are more aesthetically pleasing than the initial layout and the final layout of a given graph has the same embedding as the original graph [9]. Hong and Merrick [9] have investigated the new problem of automatic metro map layout. They have defined a set of aesthetic criteria and presented a method to produce the layout automatically. They have combined several different

layout methods. The algorithms that they have utilized are GEM algorithm, PrEd algorithm and a magnetic spring algorithm. They have added magnetic spring forces to the spring embedded model. Edges are drawn straight-line, without bends. First, each graph is simplified by removing the nodes of degree two. In their methods the following criteria are used for producing a good metro map layout:

- C1: Each line is drawn as straight as possible.
- C2: No edge crossings are present.
- C3: No overlapping of labels occurs.
- C4: Lines are mostly horizontal or vertical, with some at 45°.
- C5: Each line is drawn distinctly, with a unique color [9].

They have designed layout methods based on the first four these criteria. They have investigated five different methods:

- (1) Method 1: The GEM algorithm.
- (2) Method 2: The GEM algorithm with edge weights
 - Simplifying the metro map graph G'
 - Producing a layout L' of G' using GEM algorithm with edge weighted
 - Reinserting the removed vertices.
- (3) Method 3: Modifying PrEd algorithm with edge weights
 - Simplifying metro map graph
 - Producing an initial layout of G' using GEM algorithm
 - Producing a better layout using PrEd algorithm by edges weights
 - Reinserting the removed vertices.
- (4) Method 4: Modifying PrEd algorithm with edge weights and orthogonal magnetic spring algorithm
 - Simplifying the metro map
 - Producing an initial layout using GEM algorithm
 - Producing a better layout using PrEd algorithm by including edge weights an orthogonal magnetic field forces
 - Reinserting the removed vertices.
- (5) Method 5: Modifying PrEd algorithm with edge weights, orthogonal and 45° magnetic forces
 - Simplifying the metro map graph
 - Producing an initial layout using GEM algorithm
 - Producing a better layout using PrEd algorithm with edge weights, orthogonal and 45°magnetic forces
 - Reinserting the removed vertices.

All of this method has their advantage and disadvantage. The Hong and Merrick method do not present a good layout. In Stott and Rodgers method, there are several edges which aren't octilinear. The Nollenburg and Wolff method has a good final layout but their method is difficult to apply.

4. New Algorithms

In this section, we introduce the metro map layout problem and propose a new method for automated drawing of metro map. These methods are simple and produce the result in acceptable time.

4.1 The Metro Map Layout Problem

Let $G = (V, E, P)$, be the plane input metro map graph, consisting of a set of vertices V ; $|V|=n$, and a set of paths P . We consider that G is planar. If it is not planar, we planarize it by adding some dummy vertices. For simplification, we consider the maximum degree of vertices is 8. The main problem of this map layout is defined as follows:

Input : a metro map graph G with a set of paths.

Output: an octilinear layout of G .

First, we explain the preprocessing step. The degree of each node v in graph is calculated; according to the number of crossing edges. For each path; $P_i = \{v_i | 1 < i < n\}$; a degree of path is calculated; by calculating the sum of degree of its vertices.

$$\text{Degree } P_i = \sum_{v_t \in P_i} \text{deg } v_t \quad (2)$$

The path that has the maximum degree is called the *Maximum path* in the graph. We simplify G by removing the vertex of degree-2. These nodes do not contribute to the initial embedding. The resulting graph contains intersection nodes and nodes with degree one (extension). In our approach, the graph is embedded onto an integer square grid (fig.2.). The size of the grid is defined by g , that g is calculated by:

$$g = 2 * (\text{long of } \textit{Maximum path}) \quad (3)$$

This guarantees the sufficient space for embedding the graph in the grid. The *Maximum path* is put in the middle of grid. For each node, 8 directions are considered. The 8 directions is shown in fig.3. Two methods for embedding the graph in the grid, is proposed. In method 1, the nodes are put in grid according to the 8 directions around each node. In the method 2 each node has a weight corresponding to the number of degree-2 nodes that are removed. The temperature of the system is defined as weights. In each iteration in this method the temperature is calculated and if it is not zero the position of the weighted nodes is updated. Each time that a node is put in grid this step is reiterated. This guarantees the uniform edges.

4.2 Method 1

For each node of the *Maximum path*, all of its neighbors are put in the directions (for each node of *Maximum path* the direction 1 and 3 are already occupants). Then for the new neighbors, this step is repeated and this step is repeated until all of nodes, are put in grid. After this step, all nodes of degree two are inserted. Finally, each node has coordinated in grid and all lines are horizontally, vertically or 45°.

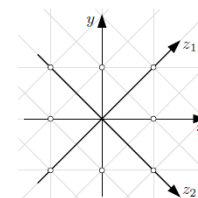


Fig.2 the grid for embedded the graph.

4.3 Method 2

Method 2 is used a preprocessing step, after removing the degree-2 nodes; consider nodes weight according to the number of the degree-2 nodes that is removed. In each iteration, after having put the nodes in grid, the temperature of graph is updated this is the sum of weights of the node in grid. If it is not zero, the position of nodes is update. There is a repulsive force between two nodes that have a weight. This repulsive force is computed and the position of nodes weighted, is changed. Their position change according to the number of nodes that are deleted. This step is repeated until all nodes are put in grid and the temperature of system is zero. The result is a graph with octilinear layout, and the length of edges is uniform.

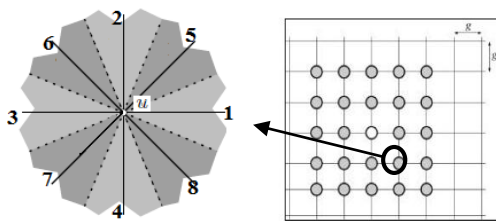


Fig.3 the 8 direction for each nod in grid.

```

Procedure Method1 is
-- Input:
-- G= (V, P) graph where
-- V= set of nodes (URL)
-- P=set of paths
-- max_path=the path with maximum weight
-- l_max_path=long of maximum path
-- g_ the size of each square in grid
-- Output: for each v∈V, a coordination
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
end;

Procedure Method2 is
-- Input:
-- G= (V, P) graph where
-- V= set of nodes (URL)
-- P=set of paths
-- max_path=the path with maximum weight
-- l_max_path=long of maximum path
-- g_ the size of each square in grid
-- Output: for each v∈V, a coordination
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
end;
    
```

Fig.4 the main loop of method1 and method2.

5. Conclusion

The Metro Map Metaphor can be used for visualising abstract information. A Metro Map can be shown as a graph. The automated method for visualisation of graphs is a challenge. In this work we have studied the problem of automatically drawing metro maps. We have proposed two methods for automated drawing metro maps. In each method, the graph is put in a grid. Then each node has a coordinate and the final graph has an octilinear layout. In second method, the graph is weighted and a temperature is used. In each iteration, the temperature decreases. The final results show that the second method has a good result for the dense graph; and the first method has satisfactory results for the graphs that have many intersections. The examples show that the results have a good octilinear layout, but in some cases there are some label overlaps.

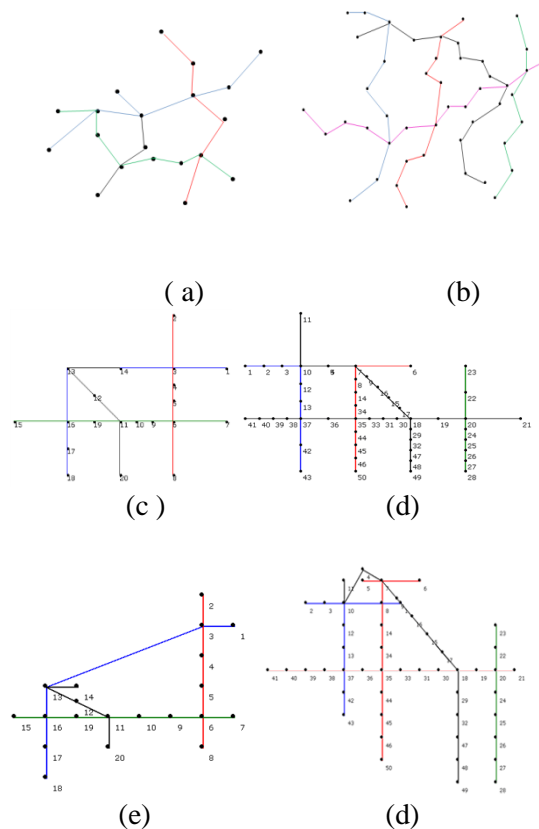


Fig.5 The result of method 1 and 2, (a)and (b) the initial graph, (c) and (d) the result of method 1, (e) and (d) the result of method 2.

References

[1] J. M. Stott and P. Rodgers ,J. C. Martínez-Ovando, S G. Walker , “Automatic Metro Map Layout Using Multicriteria Optimization”, journal of visualization and computer graphics, pp. 101-114, 2011.

[2] S. H. Hong, D. Merrick, and H.A.D. do Nascimento, "Automatic Visualisation of Metro Maps," *J. Visual Languages and Computing*, vol. 17, pp. 203-224, 2006.

- [3] T. Bang, K. Gronbak, P.S. Hansen, Using a metro map metaphor for organizing web-based learning resources, In Proc. of ED-MEDIA 2002, pp. 647–652, 2002.
- [4] K. V. Nesbitt. Getting to More Abstract Places with the Metro Map Metaphor, 2003.
- [5] E. S. Sandvad, K. Gronbak, L. Sloth and J. L. Knudsen, Metro Map Metaphor for Guided Tours on the Web: the Webwise guided Tour system, *Proc. of International Conference on World Wide Web*, pp. 326-333, 2001.
- [6] T. Bang. Knowledge Sharing in a Learning Resource Centre by Way of a Metro Map Metaphor. 67th IFLA Council and General Conference, Boston, 2001.
- [7] M. Nollenburg and A. Wolff. A Mixed Integer Program for Drawing High Quality Metro Maps. In Proc 13th International Symposium on Graph Drawing. Verlag Berlin Heidelberg, Springer, 2005.
- [8] F. Bertault, A Force-Directed Algorithm that Preserves Edge Crossing Properties, *Graph Drawing 99*, LNCS 1731, pp. 351-358, Springer Verlag, 1999.
- [9] S. Hong, D. Merrick, Automatic Visualization of Metro Maps. *Journal of Visual Languages & Computing*, vol. 17, PP. 203-224, 2006.
- [10] P. Chaleat, D. Charnay. *Programmation Html, JavaScript*, Paris, Eyrolles, 1998.
- [11] R. Davidson and D. Harel. Drawing Graphs Nicely Using Simulated Annealing. Technical Report CS89-13, Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot, Israel, ACM, 1993.
- [12] A. Frick, A. Ludwig and H. Mehldau, A Fast Adaptive Layout Algorithm for Undirected Graphs, *Proc. of Graph Drawing 94*, LNCS 894, pp. 388-403, Springer Verlag, 1995.
- [13] H. de Fraysse, J. Pach, and R. Pollack. How to Draw a Planar Graph on a Grid. *Combinatorial*, pp. 41-51, 1990.
- [14] T. M. J. Frachterman and E. M. Reingold. Graph Drawing by Force-Directed Placement. *Software Practice and Experience*, pp. 1129-1164, 1991.
- [15] S. Hong, D. Merrick, H. A. Nsacimento. The Metro Map Layout Problem. *Australasian Symposium on Information Visualization*, Christchurch, 2004.
- [16] W. He and K. Marriott. Constrained graph layout. *An international journal*, pp. 289-314, 1998.
- [17] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, 'Optimization by simulated annealing', *Science*, 220, (4598), pp. 671–680, 1983.
- [18] R. Laurini. *Information Systems for Urban Planning. A Hypermedia Cooperative Approach*, Taylor and Francis, 2001.
- [19] R. Laurini, D. Thompson. *Fundamentals of Spatial Information Systems*. Academic Press, 1993.
- [20] P. Lacomme, C. Prints, M. Sevaux. *Algorithmes de Graphes*. Ed. Eyrolles, 2th, 2003.
- [21] S. Maccari, S. Martin. *HTML & JavaScript*. Paris, Micro Application, 2004.
- [22] T. Masui, Evolutionary Learning of Graph Layout Constraints from Examples. *ACM*, November 2-4, pp. 103-108, 1994.
- [23] D. Merrick and J. Gudmundsson. Path simplification for metro map layout. In *Pro Graph drawing*, pp. 258-269, 2007.
- [24] O'Reilly book network, <http://www.oreilly.de/artikel/routemap.pdf>.
- [25] R. Tamassia. On Embedding a Graph in the Grid with the Minimum Number of Bends. *SIAM Journal of Computing*. Vol. 16. No. 3. pp. 421-444. June 1987.

Somayeh Sobati Moghadam is lecturer in Hakim Sabzevari University. She received her B.Sc. degree in Applied Mathematics in computer in 2001 from Amir Kabir University of technology and M.Sc. degree in IT in 2007 from INSA-Lyon- France. From 2010 she is lecturer in Hakim Sabzevari University. Her research interests include Information visualization, graph theory, GIS and network security.

Optimizing the Number of Neighbors in Trust Based Recommender Systems

Harmeet Kaur¹ and Deepali Jain²

¹Department of Computer Science, Hans Raj College, University of Delhi
Delhi - 110007, India

²Department of Computer Science, Hans Raj College, University of Delhi
Delhi - 110007, India

Abstract

Users in trust based recommender systems seek recommendations from their directly trusted parties known as neighbors or from neighbors of neighbors and so on. This paper proposes an appropriate range ($[\min_{rec} \dots \max_{rec}]$) providing minimum and maximum number of recommenders that one should have in his close communication. More neighbors than the maximum number of neighbors (\max_{rec}) suggested by the range does not improve the quality of recommendations but requires more time and computations in accumulation of extra recommendations. Lesser number of neighbors than \min_{rec} may compromise the quality of recommendations thus requiring seeking recommendations transitively. This again involves time and computation in propagating the query through chain of neighbors and getting the responses. A method to maintain number of neighbors within this range is also proposed. Experiments were conducted on real datasets to discover the most appropriate number of neighbors that an agent should keep.

Keywords: Trust, Recommender System, Optimal number of recommenders

1. Introduction

With the overwhelming amount of information available on the World Wide Web, it is tremendously complicated for users to pick out the best possible option for them. Information overload has become an increasingly common problem in today's large scale internet applications where users are dealing with very large amounts of data that can become time consuming to analyze [21]. Thus it is important to have tools to help users to select the relevant part of online information. A popular way to address this matter is to use recommender systems. Recommender systems are heavily used in e-commerce to provide users with high quality, personalized recommendations to help them find satisfactory items (e.g. books, movies, news, music, etc.) among a huge number of available choices[7].

In general, recommender systems suggest items by matching the attributes of an item to the profile of the user (content-based recommendation), or by correlating the profile of the user (or items selected by him) with others in the system (collaborative filtering) [23]. However, these systems do not take account of how people seek recommendations from their social networks of known individuals. Since trust is a vital ingredient of any successful interaction between individuals, among organizations and/or in society at large [2], thus trust should be incorporated in recommender systems. Trust-enhanced recommender systems refine the classical recommendation techniques, by making use of trust relationships between users in a network [5]. Trust based recommender systems provide recommendations by mining the trust network referred to as Web of Trust (WoT) among its users. These trustworthy connections among users commonly take the form of weighted trust assertions, indicating how much one user trusts another. Therefore trust based recommender systems incorporate trust network known as Web of Trust (WoT) where each user is represented by his agent and these agents collaborate on the basis of trust.

This paper determines the numeric values for \min_{rec} (minimum number of neighbors) and \max_{rec} (maximum number of neighbors) for the number of neighbors that an agent should maintain in WoT. Significance of sustaining number of neighbors from this range can be understood by considering two cases. First case is where an agent in WoT is connected to more than \max_{rec} neighbors then that agent will have extra recommendations which may include repeated and irrelevant recommendations. This not only consumes agent's time but computations also, to accumulate redundant recommendations. Second case is where an agent has less than \min_{rec} neighbors, this result in losing valuable suggestions available with other agents which again is not a good option. A solution to the

problem discussed in first case lies in removing extra number of recommenders from neighborhood where as to deal with the problem mentioned in second scenario, an agent needs to increase the number of agents in its neighborhood.

This paper also presents the procedures of increasing as well as decreasing the number of agents in one's neighborhood so as to maintain the number of neighbors in the range $[\min_{rec} \dots \max_{rec}]$. Experiments have been conducted to verify the validity of this range.

Main contributions of the paper are summarized as follows:

- (1) This paper proposes an appropriate range for the number of neighbors that an agent should maintain in its neighborhood and experiments have been conducted on real data set to demonstrate the validity of range $[\min_{rec} \dots \max_{rec}]$.
- (2) Provides a method for expansion of neighborhood through which an agent can expand its neighborhood and include some more good recommenders in its direct approach to get more and better suggestions in reduced amount of time and computations.
- (3) Presents a technique of contracting of neighborhood to reduce number of neighbors from current neighborhood of an agent and thereby reducing time and computations involved in filtering preferred recommendations from a large set of redundant recommendations.

Organization of this paper is as follows: related work is discussed in section 2. Some preliminary details of WoT are given in section 3. The proposed models of neighborhood expansion and neighborhood contraction are explained in section 4 and 5 respectively. Experimentation and results hence obtained are reported in section 6. Finally, Section 7 concludes the paper and presents some directions for future work.

2. Related Study

Trust has been extensively studied in recommender systems and successfully employed to improve classical recommendation techniques significantly. In literature, there are several algorithms for calculating trust on the web of trust network, and they use different operators and methods to infer trust.

O'Donovan and Smyth proposed a profile- and item-based recommendation that takes into consideration both the similarities among users and the trustworthiness of recommendation histories. Their trust metrics compute the percentage of correct recommendations that the user has contributed. Trust

is built up between users x and y , by measuring trust of consumer (user) on producer (recommender) y as the percentage of the correct recommendations received by x from y [16].

In [15] Golbeck introduces the Tidal Trust algorithm to estimate trust values between actor pairs in a social network. One agent infers trust rating for another by using a weighted average over all neighbors.

Massa and Avesani studied the trust-aware recommender systems [19]. Their work replaces the similarity finding process with the use of a trust metric, which is able to propagate trust over the trust network and to estimate trust weight. They propose Mole Trust which performs depth-first search, to propagate and infer trust in the trust network.

Bedi et al. in [18] proposed a trust-based recommender system for the Semantic Web; this system runs on a server with the knowledge distributed over the network in the form of ontologies, and uses the Web of trust to generate the recommendations.

Jamali and Ester [17] design the Trust-Walker approach to randomly select neighbors in the trust network formed by users and their trusted neighbors. Trust information of the selected neighbors is combined with an item-based technique to predict item ratings.

In [12] paper selection of trustworthy recommenders was done on the basis of entropy between the users. Authors have developed entropy based computational model which operates at two levels and recommenders were generated by monitoring entropy between similar users.

The model presented in [6] consists of agents, objects, and agent's profiles. In this model whenever a source agent wants to rate a particular item it asks its neighbors and its neighbors in turn pass on a query to their neighbors if they cannot provide a rating themselves. In order to generate the transitive trust from source agent to sink agent they have used the multiplicative approach and multiply the trust values along the path between the source and sink agent.

Most of the existing trust based recommender systems follow one of the following methods of assembling recommendations:

1. Getting recommendations from directly trusted associates only, i.e. only from those agents in WoT which are in the direct link of the recommendation seekers agent known as neighbors. [18]
2. Getting recommendations transitively by propagating the query through the chain of connections towards user's neighbors of neighbors and so on. [1, 3, 4, 6, 8, 9, 10, 15, 19,20, 24]

The above mentioned techniques have not considered the optimal number of recommenders while generating recommendations. However some researchers have proposed to obtain recommendations from topmost k recommenders only, but they have not provided a numeric value for k or some minimum or maximum number of recommenders. Thus this paper proposes a range providing minimum and maximum number of neighbors an agent should possess in order to retrieve useful and complete recommendations in least amount of time and effort.

3. Web of Trust

Trust based recommendation systems usually construct a trust network called Web of Trust (WoT) where nodes are users and edges represent trust between two users. It is a virtual community of agents where agents interact and cooperate with each other in order to find valuable information for their human users [11]. The goal of a trust based recommendation system is to generate personalized recommendations by aggregating the opinions of users in their trust network [22]. In WoT each agent is connected to a number of agents in web of trust which forms its neighborhood.

In WoT, trust is initialized on the basis of ability of an agent to give good recommendations and is updated using actual interactions. Boolean expressions such as trust or no trust is not appropriate for user users in social network. In real life scenarios an element of vagueness is always involved while assigning trust to a known social contact or a friend. Thus fuzzy logic is very well-suited to represent such natural language labels which represent vague intervals rather than exact values. Instead of assigning trust in crisp terms one tends to assign it in the range of 0 to 1 where 0 defines no trust and 1 symbolizes total trust. This paper uses degree of trust (where degree ranges from 0 to 1) to represent trust between two agents.

Web of Trust (WoT) can be viewed as a directed graph where:

- Agents are represented by nodes of the graph.
- Directed link from source vertex to the target vertex represents the fact that agent associated with source trusts agent linked to the target vertex.
- Weights of edges of the directed graph are annotated with the degree of trust from source to target, where this degree ranges from 0 to 1 (taking trust as fuzzy value)

Figure 1 depicts an example web of trust where nodes symbolizing agents are connected through

directed edges. Presence of the directed edge from agent a_i to agent a_j furnishes the information that a_i trusts a_j and the weight of this edge that is t_{ij} is the degree of trust from a_i to a_j which stands for the extent to which a_i trusts a_j to give good and useful recommendations.

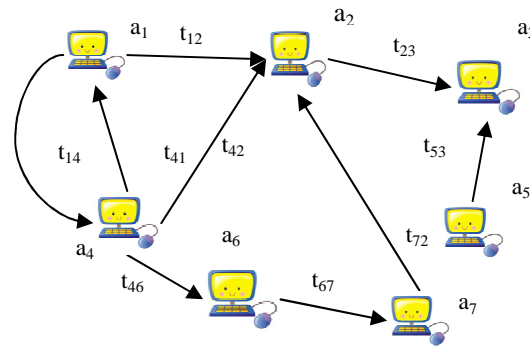


Fig. 1 Web of Trust

If there is a direct trust path in Web of Trust from agent a_i to agent a_j then agent a_i can directly take a_j 's suggestions into consideration. Here comes the significance of number of agents directly connected to an agent. If sufficient numbers of agents are in direct contact with the user agent then there is no need to propagate user query and finding trust on non adjacent agents transitively which saves time as well as computations, therefore in the situation less than required number of agents are connected to source, it can invoke process of expansion of neighborhood and include some more trustworthy agents in its direct association.

In the case where source is connected to large number of agents then again it will result in wastage of time and effort in accumulation of redundant responses, thus it has to restrict the response accumulation process by removing lesser trusted parties from its neighborhood.

There is no role of centralized authority in web of trust to maintain data repository and performing calculations to generate and process recommendations. Each agent is responsible for maintaining its data and carry out computations to generate and aggregate recommendations.

4. Expansion of neighborhood

In a scenario where source (recommendation seeker) has less than \min_{rec} number of agents in its neighborhood, it will have to propagate its request towards its neighbors of neighbors and so on until its query is satisfied. This results in involvement of time

and computations in reaching suitable and trustworthy distant agents and fetching results from those agents in addition to finding trust on those distant agents. In order to avoid the additional load of accumulating recommendations transitively each time the query propagates, an agent should increase members in its neighborhood by having good recommenders as neighbors. For the purpose of including any new agent in the neighborhood source agent will have to estimate its trustworthiness as well, thus one would require a procedure to calculate trustworthiness of newly added agents only once so that in future their recommendations could be taken without wasting much time and effort. This paper presents an algorithm of expanding neighborhood to add more agents in source's neighborhood where calculation of degree of trust for distant agents happens once and later on source can fetch their advice directly.

When source agent say a_i wishes to include some more agents in its neighborhood, it carries out the process of expansion of neighborhood. Procedure of expansion of neighborhood involves two main steps:

1. Propagation of the request from source towards its neighbors to suggest trustworthy agents that can be added into the source's neighborhood.
2. Accumulation of responses from neighbors and computing trustworthiness of newly suggested agents.

4.1 Request Propagation

As a part of request propagation procedure, the source prepares a request with the following 5-tuple query

$\langle \text{request_id}, \text{trust_threshold_neighbor}, k, \text{item_list}, \text{liking_list} \rangle$

where

- **request_id** is the unique identification number of the request,
- **trust_threshold_neighbor** defines the minimum value of trust in an immediate neighbor so that the request can be propagated to that neighbor,
- source's neighbors searches their list of acquaintances and report its k^{th} most trusted neighbor. Initially k is set to be 1. This parameter k will help in finding some more trustworthy agents in subsequent invocation of expansion process, if required, as each time agent a_i calls this process its neighbors will report different agents in decreasing order of trustworthiness,
- In order to be included in source's neighborhood, agents suggested by source's neighbors must review 'm' items provided in the **item_list** prepared by the source. These agents will be

inquired about their likes and dislikes for these items which is then used to find their similarity with the source and to judge their trustworthiness,

- **liking_list** is a list of m entries where each entry gets filled by either 1 or 0 on completion of algorithm 1, m^{th} entry of this list corresponds to m^{th} item in the **item_list**.

Source agent a_i prepares the request and finds the trust t_{ij} on all its neighboring agents a_j . For all the neighboring agents a_j such that $t_{ij} > \text{trust_threshold_neighbor}$, a_i sends a request to find trustworthy agents. When a neighboring agent a_j of a_i receives a request in the form of a 5-tuple from the source, it undertakes steps as outlined in Algorithm 1.

Algorithm 1: Request Propagation

1. a_j searches its list of neighbors
2. from all the neighboring agents of a_j , it selects its k^{th} most trusted neighbor say a_{jk}
 - 2.1 a_j retrieves and collects likes and dislikes of a_{jk} about items present in **item_list**
3. a_j populates **liking_list** of a_{jk} such that

$$\text{liking_list}[x] = \begin{cases} 1 & \text{if } a_{jk} \text{ likes } x^{\text{th}} \text{ item in the item list} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$
4. Send response as $\langle a_j, a_{jk}, \text{liking_list} \rangle$ to the sender of the request, a_i

4.2 Response Accumulation

A response is a tuple of the form $\langle \text{sender}, \text{nominee_agent}, \text{liking_list} \rangle$ where

sender is the one who is sending the response towards the source,

nominee_agent is the agent recommended by sender, and

liking_list is the populated **liking_list** of the suggested agent.

Algorithm 2 outlines the steps taken by the source when it receives all such responses from all its neighbors a_j , where every response is of the form $\langle a_j, a_{jk}, \text{liking_list} \rangle$.

Algorithm 2: Response Accumulation

1. a_i prepares $n \times 1$ column matrix N of nominee agents

$$N = \begin{bmatrix} n_1 \\ \vdots \\ n_n \end{bmatrix} \quad (2)$$

Here, n_p is the p^{th} nominee,

2. a_i prepares $n \times 1$ column matrix T

$$T = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} \quad (3)$$

Here, t_p (equivalent to t_{ij} of Algorithm 1) is trust between source and its neighbor which has recommended p^{th} nominee,

3. a_i arranges CM (choice matrix) which is the matrix of order $m \times 1$ that represents the source's likes and dislikes about the m items in the `item_list`

$$CM = \begin{bmatrix} cm_1 \\ \vdots \\ cm_m \end{bmatrix} \quad (4)$$

Here

$$cm_q = \begin{cases} 1 & \text{if source agent likes item } q \\ 0 & \text{if source agent dislikes item } q \end{cases}$$

4. DOI is also a matrix of order $m \times 1$ that represents the importance source associates with the m items in the `item_list`.

$$DOI = \begin{bmatrix} doi_1 \\ \vdots \\ doi_m \end{bmatrix} \quad (5)$$

Here $0 \leq doi_q \leq 1$ and $\sum_{q=1}^m doi_q = 1$

5. a_i carries out steps mentioned in algorithm 3 given in Section 4.3 which calculates DOT (Degree of Trust) matrix. DOT is a matrix of order $n \times 1$ having the calculated values of degree of trust on nominee agents.
6. if $DOT_p > \text{trust_threshold_new_agent}$ (6)
 - 6.1 include p^{th} nominee in `final_agents_list`
7. if number of agents in `final_agents_list` $< \text{min}_{rec}$ then
 - 7.1 $k = k + 1$ (7)
 - 7.2 source agent repeats algorithm 1 and algorithm 2.

Algorithm 1 and 2 outline the steps taken when source agent a_i wants to expand its neighborhood by adding some more trustworthy neighbors in its close association. The agent a_i asks its immediate neighbors about some credible nominee to become its neighbor. Since trust decays with the increase of number of hops along social trust path and trust decay is commonly agreed upon, for people tend to trust individuals trusted by immediate friends more than individuals trusted by friends of friends and so on [14]. Thus the process of inquiring about some trustworthy neighbors from immediate neighbors takes trust decay into account as query propagation is restricted up to one level away from source. Neighboring agents respond by giving the names of its trustworthy agent that becomes a nominee. Using algorithm 3, source computes degree of trust on the nominee agents and if this computed degree of trust is greater than `trust_threshold_new_agent`, source will include that nominee in its neighborhood. To

include more agents, if required, in the neighborhood, source will again repeat algorithms 1 and 2 with $k = k + 1$.

4.3 Computation of trustworthiness of new agents

In order to ascertain degree of trust on agents recommended by neighbors, source agent calls algorithm 2 in which it prepares nominee-item (NI) matrix of order $n \times m$, having n rows for n nominee agents and m columns for m items. Algorithm 2 populates this nominee-item matrix by accumulating nominee's likes and dislikes about m items. Corresponding to each liking for j^{th} item by i^{th} agent, a 1 is inserted in $NI[i, j]^{\text{th}}$ position of NI matrix, where as for dislike a 0 is inserted.

Algorithm 3: Computation of degree of trust on nominee agents by the source agent

1. Prepare nominee-item (NI) matrix of order $n \times m$, using matrix N and n liking_lists

$$NI = \begin{bmatrix} ni_{11} & \cdots & ni_{1m} \\ \vdots & \ddots & \vdots \\ ni_{n1} & \cdots & ni_{nm} \end{bmatrix} \quad (8)$$

Here, n is the number of nominees, m is the number of items, and

$$ni_{ij} = \begin{cases} 1 & \text{if nominee } i \text{ likes item } j \\ 0 & \text{if nominee } i \text{ dislikes item } j \end{cases}$$

2. Compute weighted NI (WNI) matrix from NI and DOI of order $n \times m$ as follows:

$$WNI = \begin{bmatrix} ni_{11} \times doi_1 & \cdots & ni_{1m} \times doi_m \\ \vdots & \ddots & \vdots \\ ni_{n1} \times doi_1 & \cdots & ni_{nm} \times doi_m \end{bmatrix} \quad (9)$$

3. Compute similarity matrix SC as product of two matrices WNI and CM resulting in SC matrix to be of order $n \times 1$.

$$SC = WNI \times CM \quad (10)$$

4. For each nominee a_p , compute the final trust on a_p as follows

$$4.1 \quad DOT = \alpha \times SC + \beta \times T \quad (11)$$

Here

$$DOT = \begin{bmatrix} dot_1 \\ \vdots \\ dot_n \end{bmatrix} \quad (12)$$

DOT is then $n \times 1$ matrix where dot_p is computed trust between source and p^{th} nominee, SC is the similarity matrix, T is the $n \times 1$ matrix, t_p is trust between source and its neighbor which has recommended p^{th} nominee,

α is the weight of similarity parameter,
 β is the weight of trust parameter, and
 $\alpha + \beta = 1$.

- Return matrix DOT as result which contains computed degree of trust on nominees

Algorithm 3 is used to compute degree of trust for agents where different weights are provided to various items. These weights are nothing but their degree of importance assigned by the source. Correlation between similarity and trust has already been proved by [4, 15, 19]. Algorithm 3 uses similarity matrix SC which furnishes the information about similarity between source and agents recommended by source's neighbors by comparing their likes and dislikes for items in item_list. Finally degree of trust from source to newly proposed agents is computed using equation (11) where parameters α and β controls relative importance of similarity index and trust value respectively. If the source is interested in looking similar agents to be included in its neighborhood then more weight should be assigned to α and on the contrary if source would like to give more weight to the path in WoT from where it is coming then $\alpha < \beta$.

5. Contraction of neighborhood

Consider the situation where source is in direct communication with a large number of agents, more than \max_{rec} , then it leads to wastage of time and effort in gathering of superfluous responses from extra neighbors, thus source has to adopt the procedure of contraction of neighborhood given in algorithm 4 to remove lesser trusted parties from its neighborhood.

Algorithm 4: Neighborhood Contraction

- a_i prepares $n \times 1$ column matrix N of neighboring agents

$$N = \begin{bmatrix} n_1 \\ \vdots \\ n_n \end{bmatrix} \quad (13)$$

Here, n_p is the p^{th} neighbor,

- a_i prepares $n \times 1$ column matrix T

$$T = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix} \quad (14)$$

Here, t_p is trust between source and its p^{th} neighbor,

- a_i retrieves and collects likes and dislikes of a_p about items present in item_list, for all p

- a_i populates liking_list of its p^{th} neighbor such that

$$\text{liking_list}[x] = \begin{cases} 1 & \text{if } a_p \text{ likes } x^{th} \text{ item in the item list} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

- a_i arranges CM which is the matrix of order $m \times 1$ that represents the source's likes and dislikes about the m items in the item_list

$$CM = \begin{bmatrix} cm_1 \\ \vdots \\ cm_m \end{bmatrix} \quad (16)$$

$$\text{Here } cm_q = \begin{cases} 1 & \text{if source agent likes item } q \\ 0 & \text{if source agent dislikes item } q \end{cases}$$

- DOI is also a matrix of order $m \times 1$ that represents the importance source associates with the m items in the item_list.

$$DOI = \begin{bmatrix} doi_1 \\ \vdots \\ doi_m \end{bmatrix} \quad (17)$$

$$\text{Here } 0 \leq doi_q \leq 1 \text{ and } \sum_{q=1}^m doi_q = 1$$

- a_i carries out steps mentioned in algorithm 3 given in Section 4.3 and obtains DOT (degree of trust) matrix where DOT contains new degree of trust from source to its neighbors

- a_i arranges all its neighbors in descending order on the basis of DOT matrix.

8.1 Source keeps x most trusted neighbors from this list, where x lies in the range $[\min_{rec} \dots \max_{rec}]$ and remove others from its neighborhood.

For the purpose of removing some agents from its neighborhood, source carries out algorithm 4. Source agent judges all its neighboring agents based on the similarity between itself and its existing neighbors and their current degree of trust by utilizing algorithm 3 (here instead of nominees, neighbor's information is utilized) and acquire matrix DOT, which is the matrix having updated degree of trust on existing neighbors. The source then sorts the list of neighbors, in descending order of their newly obtained trust stored in matrix DOT and keeps x most trusted agents in its neighborhood, where $\min_{rec} < k < \max_{rec}$. Algorithm 4 serves dual purpose of contraction of neighborhood and pruning of neighborhood on the basis of changes in source's taste of items. Without application of algorithm 3 simply removing less trustworthy agents from neighborhood do not take source and neighbor similarity into account. In this manner source can maintain number of neighbors from suggested range by applying algorithms 1 to 4.

6. Experimental Setup

Experiments were carried out to determine the appropriate range $[\min_{rec} \dots \max_{rec}]$ of neighbors. The dataset for experiments was derived from web community of Apartmentratings.com. The data set rates thousands of apartments in USA on the seven criteria viz. Parking, Maintenance, Construction, Noise, Grounds, Safety and Office Staff. The above set of parameters describes basic features of an apartment, according to which recommender will describe the apartment and probable user will choose the apartment to live in. For experiments the data has been collected directly from the Apartmentratings Web site [13]. The dataset consists of approximately 500,000 raters who rated a total of almost 1000 different apartments at least once. The total numbers of reviews are around 1,000,000. Out of 500,000 raters, 10 different sets of 50 raters were chosen as sample to study algorithms. Thus in total 500 raters were chosen. For each set of 50 raters their corresponding 50 agents were created using JADE and profile of each user was placed in its agent. The system is implemented using Java and JADE platforms. Algorithm of recommendation generation and algorithms 1 to 4 of expansion of neighborhood and contraction of neighborhood are developed and implemented as Java classes and are integrated with the JADE platform. The interaction among different agents for developing trust relationships were implemented as agent behaviors. In the initial phase of the experiment for each of its 50 users their list of acquaintances along with the degree of trust that they can place on each other were generated randomly. According to these lists initial web of trust was spawned which is similar to web of trust in figure1 but with 50 agents. Web of trust thus contains a directed edge from an agent to all the agents in its list of acquaintances weighted by the degree of trust as reported in the randomly generated list and hence become its neighbors. This was done for each set of 50 agents.

6.1 Discussion

To discover the appropriate range of minimum and maximum number of neighbors that an agent should possess, simulations were carried out with for each set of 50 agents by making each agent as source. In each simulation source agent initiated the process of recommendation generation with different number of neighbors in its neighborhood. For the purpose of estimation of range $[\min_{rec} \dots \max_{rec}]$ simulations were carried out by having one neighbor and extended up to 15 neighbors in source neighborhood. For each set of 50 agents experiments were carried

out and their results were documented independently as well as their average result was also recorded. Figure 2 illustrates data obtained after running experiment for the first agent from first set of 50 agents.

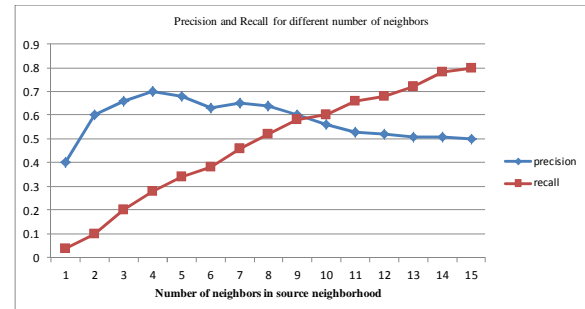


Fig. 2 Precision and recall obtained for a_1 for different number of neighbors

The two metrics commonly used to evaluate the recommender systems are precision and recall.

Precision is defined as the fraction of the selected items that are relevant to the user's needs. It measures the selection effectiveness of the system and represents the probability that the item is relevant.

$$\text{Precision} = \frac{\text{Number of relevant recommendations retrieved}}{\text{Total number of recommendations retrieved}} \quad (18)$$

Recall is defined as the ratio of the relevant items selected to the total number of relevant items available. Recall represents the probability that a relevant item will be selected.

$$\text{Recall} = \frac{\text{Number of relevant recommendations retrieved}}{\text{Total number of relevant recommendations available}} \quad (19)$$

Considering equal error rate or equal accuracy which denotes the intersection of precision and recall curves, from the figure 2 it can be clearly seen that at the point of having 9 neighbors in a_1 's neighborhood, precision and recall intersect. This graph advocates that an agent must possess 9 neighbors in its neighborhood. Similarly other runs were also carried out for first 50 agents and their readings were recorded. Another such set of 50 agents was taken to determine suitable range of neighbors and figure 3 demonstrates the result.

In the figure 3 the intersection point of precision and recall comes out to be 13. Thus according to figure 3 agents should maintain 13 neighbors in their neighborhood.

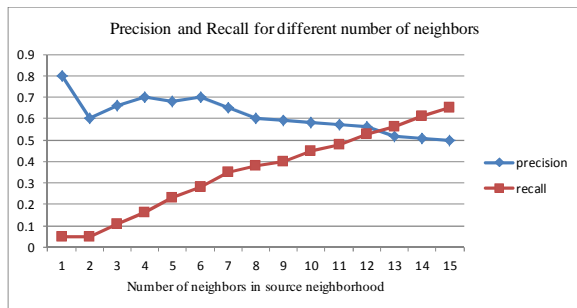


Fig. 3 Precision and recall obtained for different number of neighbors

Figure 4 demonstrates the result obtained after taking average of all the simulation of 10 sets of 50 agents in each set.

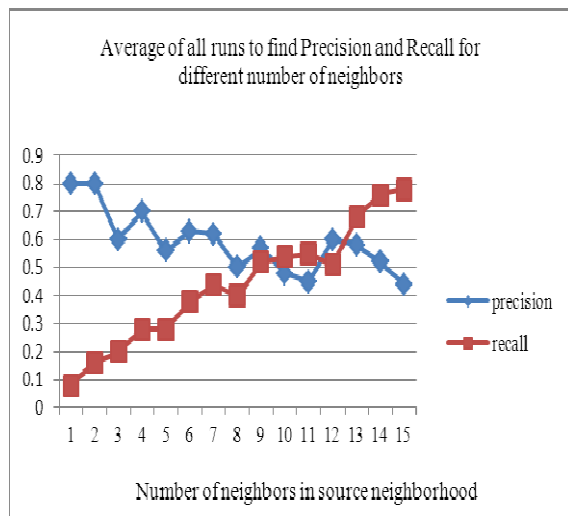


Fig. 4 Average precision and recall for all simulations

It is clearly evident from figure 4 that an agent in WoT should have between 9 to 13 neighbors which defines the optimal range for the number of neighbors that an agent should maintain. This range can be inferred from the graph as precision and recall intersect thrice: first at 9, then at 11 and again at 13. Hence, the two extreme intersections give the range on number of neighbors as [9 ...13].

7. Conclusion and Future Work

In this paper a suitable range of minimum and maximum number of neighbors that an agent should preserve is proposed. In order to maintain number of neighbors from this range two procedures are presented. Firstly, a process of expansion of neighborhood to assist an agent to enhance its set of recommendations by discovering good recommenders and make them its trustworthy

neighbors is proposed. Secondly, a procedure of contraction of neighborhood to eliminate lesser trusted neighbors from agent's neighborhood is proposed. Proposed techniques involve discovering similarity between source agent and other agents and computation of trust on the basis of similarity factor and initial trust. Our process of expansion of neighborhood takes trust decay into account as query propagation is restricted up to one level away from source. Experimental results have demonstrated that the proposed range follows the real life pattern where one tends to have a sufficient number of close communicates, very few friends provide limited knowledge where as more than required inundate one's database.

More experiments are being conducted with some other real social network datasets to further validate the results. A feasible remedy for the situation where the source needs to enlarge its neighborhood and its direct neighbors are not able to provide adequate number of trustworthy candidates is under consideration.

References

- [1] A. Jøsang, E. Gary, and M. Kinateder, "Analysing topologies of transitive trust", in 1st International Workshop on Formal Aspects in Security and Trust, 2003.
- [2] A. Medic, and A. Golubovic, "P2P Appliance Calculation Method For Trust between Nodes within a P2P Network", International Journal of Computer Science Issues, Vol. 9, No. 4, No 2, 2012, pp 125-130.
- [3] B. Alcalde, and S. Mauw, "An algebra for trust dilution and trust fusion", Formal Aspects in Security and Trust, 5983, Springer, 2009, pp. 4-20.
- [4] C. Ziegler, "Towards Decentralized Recommender Systems", Ph.D. thesis, University of Freiburg, 2005.
- [5] D. Doherty, S. Jouili, P. V. Roy, "Trust-Based Recommendation: an Empirical Analysis", in 6th SNA-KDD Workshop '12, 2012, Beijing, China.
- [6] F. Walter, S. Battiston, and F. Schweitzer, "A model of a trust-based recommendation system on a social network", Autonomous Agents and Multi-Agent Systems, Vol. 16, No. 1, 2008, pp. 57-74.
- [7] G. Guo, J. Zhang, and D. Thalmann, "A Simple but Effective Method to Incorporate Trusted Neighbors in Recommender Systems", In 20th International Conference on User Modeling, Adaptation and Personalization (UMAP), 2012, pp 114-125.
- [8] G. Liu, Y. Wang, and M.A. Orgun and E.P. Lim, "A heuristic algorithm for service provider selection in complex social networks", in SCC'10, 2010, pp. 130-137.
- [9] G. Liu, Y. Wang, and M.A. Orgun and E.P. Lim, Trust Transitivity in Complex Social Networks, in

- Twenty-Fifth AAAI Conference on Artificial Intelligence, 2011.
- [10] G. Liu, Y. Wang, and M.A. Orgun, "Optimal social trust path selection in complex social networks", in AAAI, 2010, pp. 1397–1398.
- [11] H. Kaur, and D. Jain, "Modulating Friendship Trust into Context Level Trust in Trust Based Recommender Systems", International Journal of Recent Trends in Engineering, Vol 3, No. 2, May 2010, pp. 58-63.
- [12] H. Mehta, SK. Bhatia, P. Bedi, and VS. Dixit, "Collaborative Personalized Web Recommender System using Entropy based Similarity Measure", International Journal of Computer Science Issues, Vol. 8, Issue 6, No 3, November 2011, pp 231-240.
- [13] <http://www.apartmentratings.com>.
- [14] J. Gimpel, K. Karnes, J. Mctague, and S. Pearson-Merkowitz, "Distance-decay in the political geography of friends-and-neighbors voting", Political Geography, Vol. 27, 2008, pp. 231–252.
- [15] J. Golbeck, "Computing and Applying Trust in Web-Based Social Networks", Ph.D. Thesis, University of Maryland, 2005.
- [16] J. O'Donovan, and B. Smyth, "Trust in recommender systems", In International Conference on Intelligent User Interfaces, San Diego, California, January 2005, pp. 167–174.
- [17] M. Jamali, and M. Ester, "Trust Walker: a random walk model for combining trust-based and item-based recommendation", in 15th ACM SIGKDD international conference on Knowledge discovery and data mining, 2009, pp. 397-406.
- [18] P. Bedi, H. Kaur, and S. Marwaha, "Trust based Recommender System for the Semantic Web", in 20th International Joint Conference on Artificial Intelligence (IJCAI '07), Hyderabad, India, 2007.
- [19] P. Massa, and P. Avesani, "Trust-aware recommender systems", in RecSys'07, USA, 2007.
- [20] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust", in WWW'04, 2004, pp. 403–412.
- [21] S.MAGUREANU, "Distributed Trust-Aware Recommender Systems", Master's thesis, ICT, 2012.
- [22] S. Nepal, C. Paris, and S. K. Bista, "SRec: a social behaviour based recommender for online communities", 3rd Workshop on Social Recommender System at UMAP, 2012.
- [23] T. Heath, E. Motta, and M. Petre, "Person to Person Trust Factors in Word of Mouth Recommendation", in CHI2006 Workshop on Reinventing Trust, Collaboration, and Compliance in Social Systems (Reinvent06), 2006, pp. 22-27.
- [24] U. Kuter, and J. Golbeck, "Sunny: A new algorithm for trust inference in social networks using probabilistic confidence model", in AAAI'07, 2007 pp. 1377–1382.

Dr. Harmeet Kaur received her Ph.D. in Computer Science from the Department of Computer Science, University of Delhi, Delhi, India in 2007. She is an Associate Professor in the Department of Computer Science, Hans Raj College, University of Delhi. She has about 15 years of teaching and research experience and has published more than 20 research papers in National/International Journals/Conferences. Her research interests include Multi-agent Systems, Intelligent Information Retrieval Systems, Trust and Personalization.

Ms. Deepali Jain is a Research Scholar and working as an Assistant Professor in the Department Of Computer Science, Hans Raj College, University of Delhi. Her research area is Knowledge Based Systems and is currently pursuing PhD under Dr. Harmeet Kaur from Department of Computer Science, University of Delhi.

The Difference Between Dwelling And Home In Architecture

SHADAB SHIDFAR

College of Architecture And Urban Design, Iran University Of Science And Technology
Toronto, Ontario, M2N7B2, Canada

Abstract

In the low rating city environment which is the accomplishment of a society that makes technology rule over anything, home and dwelling have become the last shelters for human being. Considering today's life demands (Speed/Crowdedness/Noise Pollution/Technological improvements and fading communication at home) necessity of attention towards home as a shelter for tranquillity, stillness and a space for gathering and presence of all family members after a hard day's work, gets more important. Among all spaces around man, home is the most direct space related to man that affects him and gets affected by him on a daily basis.[1] It is the first space that man experiences a feeling of belonging. All our five senses are continually surfing it and are adopted to it as soon as possible. This article has a quick review on the concepts of dwelling and home, their role in human's life, and it insists on analyzing and separating these two concepts. Although these two concepts totally differ from each other, but usually are considered the same. It also states the necessity of difference between dwelling and home in relation to human being, by proposing arguments about human's relation to home and dwelling and the importance of understanding this relation for designing (specially architecture and urbanization) to create a framework to analyze the relation between human and environment.

Keywords: Dwelling, Home, Living Environment, Human

1. Introduction

Today, fundamentals of family in all societies are gradually being broken and the role of family as a society of a few people is fading and family as a one person society is getting more concrete. In this situation, everybody is about to develop and grow the quantities of their life. Now, human beings are invited to a competition. The competition is about getting more benefit and developing life standards. With technology development, human needs are getting extended more and more and older productions are being replaced by modern technologic productions. In this circle, everybody is attempting and trying to be up-to-date, without even thinking that these attempts and efforts are never going to end.

Now, to us, even house has become an industrial production like any other. We use the term "building industry" in our daily conversations and as a result, it is obvious that as time goes by, we have to destroy the old house and build a new one. New cell phone, new computer, new car and maybe a newer house. This is the utmost ideal of human, who has lost his essence and spirit.

In the low rating city environment which is the accomplishment of a society that makes technology to rule over anything, personal house has become the last shelter for human being. Therefore, personal house in the small world, is a social concept against the old personal philosophy. This, itself, is an opportunity to understand that how architecture should be related to life environment, specially houses. Excessive spaces available in house design can be a starting point for designing all kinds of different architectural spaces in different houses. At first, it seems clear that by using the terms "dwelling" and "home", we mean the same meaning and concept for both, and in a glance, there is no necessity to define these two words, because both of them are a place to provide comfort and convenience for human beings. But with a little thinking, regarding the point that the concept of dwelling is something more than a physical place, we find out that it is totally different from the concept of home.

In this article, a little bit of qualitative and quantitative values of dwelling and home as valuable spaces in man's personal life, is proposed and reviewed. So, we go through with the definition of dwelling and home and also their situation in people's private life.

2. Definition of Dwelling and its concept

From a literal point of view, in Persian language, Dwelling is equal to the word "dwelling", which means a place for getting tranquility and residence, and it is called to a place which a human lives in it. In Persian Dictionary, "dwelling" is defined: A place for settling and being, home, a lake of residence, a still place.[2]

Heidegger argues that the manner in which we dwell is the manner in which we are, we exist, on the face of the earth - an extension of our identity, of who we are. In his article "building, dwelling, thinking", he opens up old indications

and conceptual areas on the root of the German word "Bauen" (to build) to explain a soul's ailments and interprets dwelling as being with things: as he debates the residence of human essence in the truth of being, in being and time. He insists that thinking and poetry are the necessities for dwelling. He assumes dwelling as the most fundamental characteristic regarding mortals. At the end of his article we read: "On all sides, we hear talk about the housing shortage, and with good reason... What if man's homelessness consisted in this, that man still does not even think of the proper plight of dwelling as the plight?". Heidegger answers the question as follows: "This they accomplish when they build out of dwelling, and think for the sake of dwelling".[3]

On the Second UN Conference on Human Settlement in 1996 which was held in Istanbul, Turkey, a proper dwelling is defined as follows: "A proper shelter is not only the existence of a roof over people's head. A proper shelter means proper comfort, a proper space, suitable physical approach and security, ownership security, structural stability, lighting, ventilation, proper heating system, primary substructure such as water, health-care and education, garbage evacuation, proper environmental elements, proper health factors, accessible to work place and accommodations and all these factors must be provided according to financial status of people".[4]

What seems obvious at first is that the meaning of dwelling is clear and there is no need to define it, but with a little thinking about the point that the definition of dwelling is something more than just a physical place, we understand that we have to have more focus on its definition. Dwelling is a word which seems to have a simple meaning but actually consists of a wide range of definitions. And it is also a very wide and complex affair. The importance of dwelling comes from the point that many primary needs of human such as preparing food, rest, etc. is provided in it. The concept of dwelling, other than a physical place, includes the whole residence including all necessary facilities and services for well-being of the family, and also working, education and health-care programs for people.[5]

From what Schulz says, we understand that discovering humane identification is in making meaningful relations with a world comprised of things which are depended on material form of a place as if he states that the double aspects of dwelling are identity discovery and locality determination. This double aspects of architecture in buildings, are coincided on the architectural functions of visualization and admission, thus every place in its state of providing possibilities to other functions accepts them inside itself, and also takes new concepts and meaning with them.

Heidegger in his article "building, dwelling, thinking" states that getting an identity is possible through dwelling

and is in relation with understanding things and on this basis, has a phenomenological approach towards the universe and location for identification discovery, as he believes understanding phenomenon (specially location) plays a basic role in identity retrieval and the essence of human.[6]

Regarding the explanations about the concept of dwelling in last paragraphs, we can have a conclusion as the following: In fact, common definition and concept of dwelling is not a residential unit but the whole residence of living (Pictures 1 & 2). A simple definition states that a residential unit is a space consisting of at least one unit of kitchen, a room, and a private toilet, which constitutes the whole or a part of a residential building and has private exit to a street or an alley... or to a public stairway or corridor.



Picture No(1):Habitat by Moshe Safdie

A residential building is a space limited, fixed and independent from its neighbor grounds and buildings, that all or part of it is roofed and is built for the intention of living and has no special space for economic activities (service or producing) and is traded as a place for living.



Picture No(2): A residential building is a space limited

3. Definition of home and its concept:

To study the definition of home and its concept, first we refer to some quotations of some great architects about home:

Mr. Pirniya: Home is a place that its occupants do not feel uncomfortable, and the lady of the house and her children who spend their time in it, should not feel tired of it.[7]

Charles Moore: Home is the center of the world to its occupants. (To a residential area, the most indicative kind of building to enforce its locality).

Amos Rapoport: Home, in the first place, is an institution, not a structure for complex matters. For cultural matters, it is under the influence of culture. Even at first, the concept of home has not been a totally functional space. The positive point of creating an ideal environment for the family as a social environment points to home as shelter, but the important thing is that this makes sheltering as an obligation for this social institution. From the first day, man created his home because of his beliefs in rituals and cultural issues.[8]

Gaston Bachelard: Home, before anything, is an inner space and location. Its inner space gets a meaning in relation to its outside, alley and street, river and farm, village and city, sea and earth and sky, nature and world, and at last the whole universe. Its most important virtue is to raise dreams. And finally, home provides the facility, for us, to drown into our dreams, in tranquillity.[9]

We can conclude that: Home, being posed as an architectural structure in the environment, defines our identity and provides security for us, and finally we find relaxation when we set foot in it. It's at home that we find things that we know and cherish them. We bring them

from outside and live beside them because they form a part of our world. When realizing a private residence, we try what is called "Internal Relaxation"

So the concept of home in man's life is not a general definition, but a united purpose which we live in it and it preserves us from the outside world.

4. Conclusion: (An analysis on the difference between Home and Dwelling in Iran)

In Iran, home gets its real meaning beside a more complex concept, Family. Family is the most basic social institution which in Iran, an Eastern country, has an introverted function. Home and family are two inseparable elements. As family has a non-stop functionality, and is unchangeable, home also, has a lasting and unchangeable concept.

In other words, for an Eastern man, home is not like a hotel or a temporary habitation. Home is not an old car that we can change it into a new one as we reach better economic situation, or is not a temporary habitant that when we get more earnings, add to its extent. Home has a deep meaning, on the same level as family.

In Iran's classic architecture, when we look at Tabatabaees (Picture 3) or Borujerdis (Picture 4) dwellings, we always bring the family's name before their place of residence. Never the great Borujerdi built that brilliant dwelling with the intension of temporary residence, because if that was so, they would've changed their house many times and their name wouldn't last long.



Picture No(3):Tabatabaees House,Iran



Picture No(4): Borujerdis house,Iran

According to our research, we can have its result as follows: If we want to compare the words "dwelling" and "home" from a literal point of view, the definition of home is not just equal to a place to reside. In Persian language, some words receiving the suffix "KHANEH" meaning "home" get a wide range of meanings and dimensions which are related to "Residing". Residing creates cities with their culture. So we can say that home is almost a subsidiary of dwelling.

As a result we imagine home as a three dimensional space which surrounds a family of definite number of people. Home is the demonstrator of a fixed relationship between a more personal environment compared to other fixed spaces. Also the extent of home is limited to the space which man can reach everything, the architecture and whatever that mixed in it and man connects to it (Micro scale). So, home cannot be a place for different characteristics (because in a wide extent, political, economic and social situations have a background and cannot be disregarded, but home is a totally private place to secure the man from other things).

As Schulz says, residence can express the general situation and identity discovery. Comparing above subjects, we conclude that residence can express the creation of a meaningful relation between man and environment around him that this relation is a result of being identified, meaning to have feeling of being related to a place.

Dwelling is a collection of all the spaces which provide different needs of human such as proper relaxation, proper space, physical approach and proper security, ownership security, stability of the structure, lighting, ventilation, proper heating system, primary facilities like water, health-care, and education, garbage evacuation, proper quality of environment, suitable place, etc. (Mass scale).

Also, among the spaces around, maybe it can be said that home is a place that man feels a personal belonging to it. Home, being posed in an architectural framework, saves man from outside space and brings privacy to him and at last, when man sets foot in it, provides a totally personal

space for him. Home is the first place which man's very first experiences happen in it. Being alone, or with your wife, children and others, all of it is possible at home, without any external disturbance. Here, the definition of home is totally against the definition of dwelling, because our perspective of dwelling, other than the physical place, is the whole residential place and so a public place consisting of all necessary facilities and services for family living. In other words, dwelling is a more general concept compared to home, and the first concept which comes to mind after hearing the word "residence". Dwelling is residing in a place (a place where residence has different kinds: natural, complex, public and private), a space consisting of some buildings which along with primary needs, provide other needs of human being, regarding their necessity.

Home is a three dimensional limited space. Although we can partly break these limitations with correct designing, but maybe we can say that home is limited to its own definition and that dwelling has a more extended functionality than home.

References

- [1] Ross C. Brownson, PhD, Christine M. Hoehner, PhD, MSPH, Kristen Day, PhD, Ann Forsyth, PhD, James F. Sallis, PhD, **Measuring the Built Environment for Physical Activity State of the Science**, American Journal of Preventive Medicine,36(2009),p.p.99-123
- [2] Moin,Dr Mohammad Moin,**Persian Dictionary**,Amirkabir Publisher,1996,P,P,1900-1901
- [3] Hayeri,Mohammadreza,**Concept of the home**,Naghshe jahan publisher2010.
- [4] Nurberg,Schiltz,Christine,**Hidegers Thinking in architecture** , Memar Publisher,No 12,2002.
- [5] Nurberg,Schiltz,Christine,**Hidegers Thinking in architecture** , Memar Publisher,No 1,2002.

[6] Pirniya,Phd Mohammad karim,**Islamic Architecture,**
Keyhan publisher,No 8,P,PP29-40,1994

[7] Rapaport,Amos,**Humaneterian home,** Herfe honarmand
publisher,2009.

[8] Mozayeni,Manoochehr,Being and architecture,Naghshe
jahan publisher,2005

First Author Mrs.Shadab Shidfar.MS Student in Iran University Of
Science And Technology. college of architecture and urban
design.

A Novel Design of Ultra-Wideband Antenna (UWB) Using Slot-Based Structures

MOHAMED TAOUZARI¹, AHMED MOUHSEN¹, JAMAL EL AOUI¹, JAMAL ZBITOU²

¹ Laboratory MMH FST Settat University Hassan 1
Settat, Morocco

² Laboratory LITEN, Faculty polydisciplinary
Khouribga, Morocco

Abstract

Radio communications using ultra wideband (UWB) frequency have rapidly developed since the beginning of the year 2000. This technique offers advantages that made it attractive to a large number of civilian and military applications. The aim of this work is the study of the geometry features of a UWB antenna and the improvement of the bandwidth by inserting slots. To obtain a very large bandwidth operating from 2 GHz to 11 GHz we have inserted slots both on ground plane and antenna.

Keywords: planar antenna, Ultra Wide Band Antenna, Slots.

1. Introduction

The printed antennas have been involved for many applications in telecommunication systems. Modeling of several parameters in antenna's structures promotes higher performances.

The technical advantages of printed antennas converge for many success results such as:

The low weight, the minimal thickness congestion, and the possibility to generate a circular polarization adapted to curved surfaces.

In order to qualify antenna and integrate it into different systems, some disadvantages must be avoided, for instance: low bandwidth, small gain and strong influence of dielectric substrate.

In this paper, we are interested to study the design of patch antenna using Ultra Wide Band (UWB), numerical results

of slot-based structures are presented to illustrate the behavior of reflection coefficient and radiation pattern.

2. Design of UWB antenna

The conception relies on transmission line method. The below list of equations obtained from the transmission line matrix (TLM) can be used to calculate effective length and width of the patch antenna [1]. The structure of the designed antenna is shown in Fig.1.

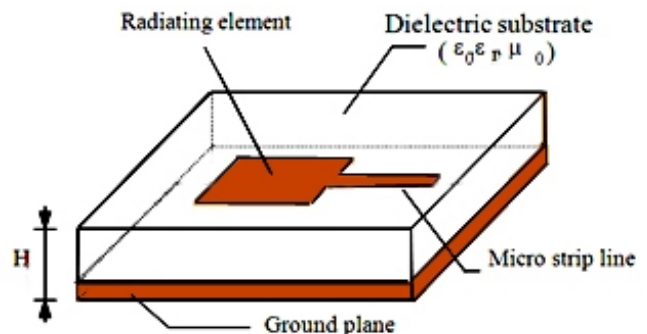


Fig 1: Structure of the patch antenna.

The dimensions of the patch antenna are given as follows [2]:

$$W = \frac{c}{2 \cdot f_r} \cdot \sqrt{\frac{2}{\epsilon_r + 1}} \quad (1)$$

$$f_r = f_0 \cdot \left[\frac{1 - \frac{2 \cdot h}{\epsilon_{eff} \cdot l \cdot \pi \cdot \alpha}}{1 + \frac{2 \cdot h}{\epsilon_{eff} \cdot l \cdot \pi \cdot \alpha} \cdot \ln \left(\frac{\sqrt{\epsilon_{eff}} \cdot 2 \cdot l}{\gamma \cdot h} \right)} \right] \quad (2)$$

$$f_0 = \frac{c}{2 \cdot l \cdot \sqrt{\epsilon_{eff}}} \quad (3)$$

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \cdot \left(1 + \frac{10 \cdot h}{w} \right)^{-1/2} \quad (4)$$

$$L = L_{eff} - 2\Delta L \quad (5)$$

$$L_{eff} = \frac{c}{2f_0 \sqrt{\epsilon_{eff}}} \quad (6)$$

$$\Delta L = 0.412 h \frac{(\epsilon_{eff} + 0.3) \left(\frac{W}{h} + 0.24 \right)}{(\epsilon_{eff} - 0.258) \left(\frac{W}{h} + 0.8 \right)} \quad (7)$$

For this antenna, the dielectric substrate used is FR4 with:

Dielectric permittivity $\epsilon_r=4.3$

Substrate height: $h=1.58\text{mm}$

Dielectric losses, $\tan\delta=0.025$

Thickness of metallization: $t=35\mu\text{m}$

f_r = radiation frequency.

$\gamma = 1.78107$

$c = 3 \cdot 10^8$ en m/s

h = dielectric height (1.58 mm of FR4)

w = width of the patch

L = length of the patch

ϵ_r = dielectric constant

3. Simulation and measurement:

Firstly, we simulated the patch antenna without slots. In the second step, we have done a new simulation taking into account including slots at the front of antenna and finally we did a third simulation by adding the slots at the ground plane [3].

3.1 Antenna structure without slot

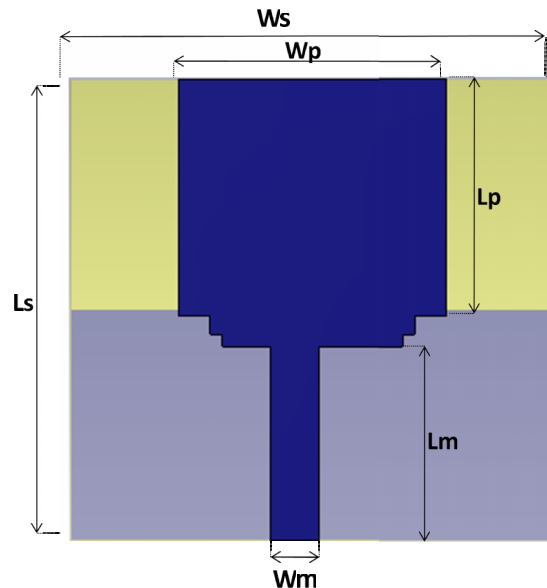


Fig 2: antenna structure without slot.

Table 1: Dimension by (mm) patch antenna.

Ws	Ls	Wp	Lp	R	Wm	Lm
40	38	22.22	19.5	7	4	15.6

The simulation of the structure is illustrated in fig 2.

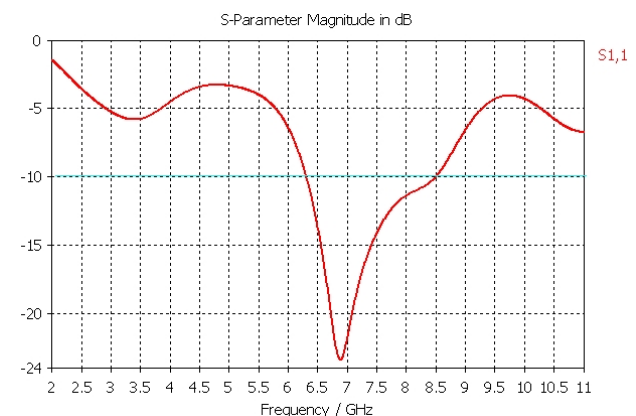


Fig. 3: Parameter S11 versus frequency.

3.2 Antenna structure with slots

In order to improve the bandwidth of the rectangular patch antenna, we have introduced slots into the antenna. The

results are in good agreement with those done by Duroc et al [4]:

3.2.1 Slots at the antenna's front

In this part, we have studied the influence of the slots on the antenna (fig. 4) and the results are shown in fig.5.

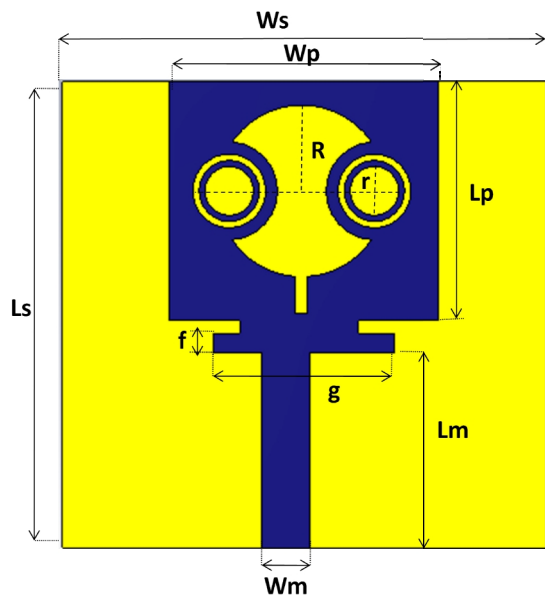


Fig 4: Antenna structure with slot.

Table 2: Dimensions in (mm) of patch antenna

g	F	R	r
15	1.6	7	2

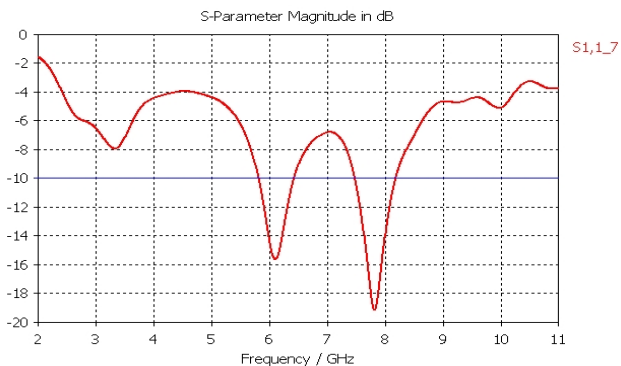


Fig. 5: Parameter S11 of the patch antenna with slots on the upper layer.

Fig.5 shows the effect of the slots on the increase of bandwidth, and we extend the insertion slots on the ground plane.

3.2.2. Slots at the ground plane

It is worth mentioning that the presence of an opening in the ground plane of a rectangular micro band patch antenna inevitably affects the resonance properties of the antenna (fig.6), it has been proved, for the micro band antennas, that this opening can be used as a way to adjust the resonance frequencies. So the point is that this opening in the ground plane adds new conception parameters that can be used as adjustments components [5].

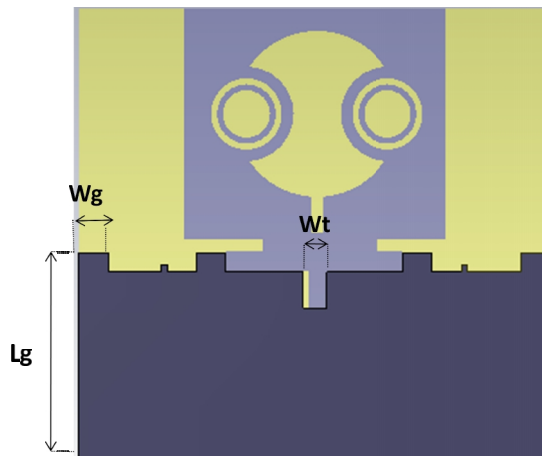


Fig 6: Slots in antenna and in ground plane.

Table 3: Dimensions in (mm).of ground plane

Lg	Wg	Wt
17.4	2.5	2

According to fig 7, we notice that we could have a band extending from 2.3GHz to 11.6 GHz thus an extension around 7 GHz compared to the first conception, hence the importance of inserting slots.

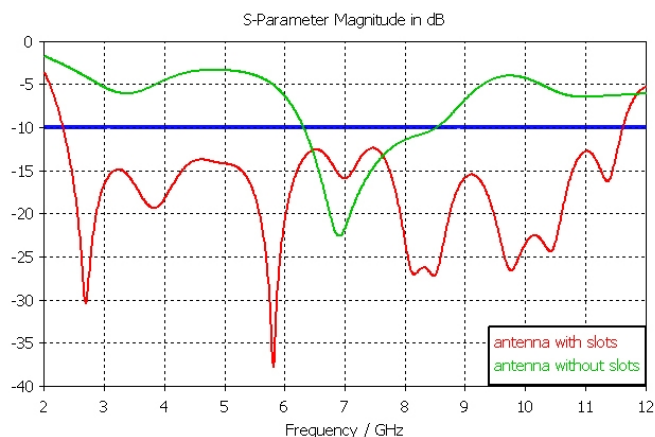


Fig. 7: Development of bandwidth.

The radiation pattern of the simulated antenna is described in fig 8; it represents a Gain of 4.94dBi for 5.8 GHz frequency and an opening angle of 89 degrees.

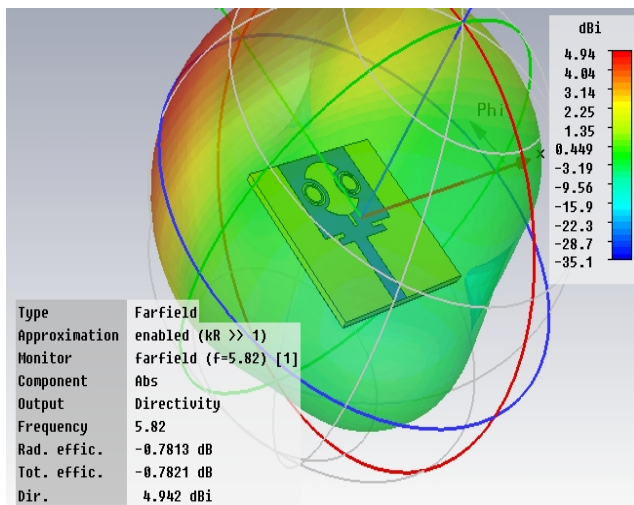


Fig.8: Radiation Pattern at f=5.8GHz.

Fig. 9 shows the simulated maximum gain of the proposed antenna. Note that the gain drops significantly at the notched bands. However, the proposed antenna shows good gain performance at other frequencies across the UWB band.

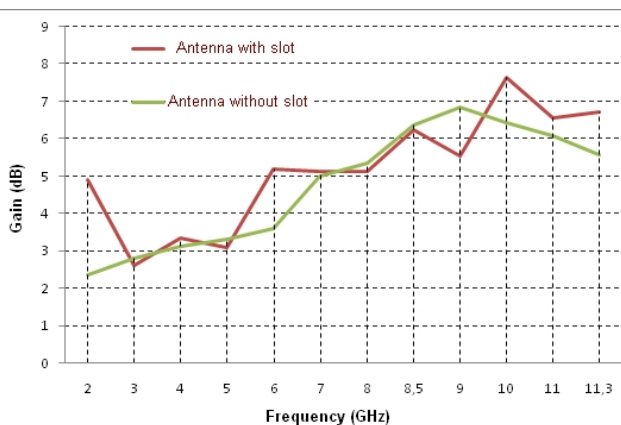


Fig.9: simulated gain are plotted here as a function of frequency.

4. Conclusion

Throughout the study and the conception of a planar antenna with a bandwidth of 2 GHz. By inserting slots at the front of the antenna and on ground plane we notice a significant improvement of this bandwidth. Thus is easy to use it in micro-wave circuits.

References

- [1] C. A. Balanis, "Antenna Theory: Analysis and Design", John Wiley & Sons, second edition, 1997.
- [2] D. M. Pozar, "Microwave Engineering", John Wiley & Sons, second edition, 1998.
- [3] Y. Duroc, A. Ghiotto, T.P. Vuong, S. Tedjini, "UWB Antennas: Systems with Transfer Function and Impulse Response", IEEE, Communication in Transactions on Antennas Propagation, Vol. 55, Issue 5, pp. 1449-1451, May 2007.
- [4] Y. Duroc, A. Ghiotto, T.P. Vuong, S. Tedjini, "On the Modeling of Ultra-Wideband antennas", IEEE Transactions on Antennas and Propagation, Vol. 55, Issue 11, Pages 3103 - 3105, November 2007.
- [5] Y. Duroc, A. Ghiotto, T.P. Vuong, S. Tedjini, "On the Characterization of UWB Antennas", Wiley, International Journal of RF and Microwave Computer-Aided.

Criticality estimation of IT business functions with the Business Continuity Testing Points method for implementing effective recovery exercises of crisis scenarios

Athanasios Podaras¹ and Tomas Žizka²

^{1,2} Department of Informatics, Technical University of Liberec, Liberec, Czech Republic

Abstract

The primary goal of the present paper is the introduction of a new approach of defining IT unit business functions' exact criticality levels and respectively categorize them to the appropriate recovery tests, prior to their thorough documentation which includes actual desired recovery timeframes. The method is entitled as Business Continuity Testing Points and it is based on the concept of Use Case Points, a fundamental project estimation tool utilized for sizing of object-oriented system development. The aim of the contribution is to ameliorate the existing manual way of determining recovery time of IT business functions that is based exclusively on experience of IT personnel, by introducing a calculation method of multiple factors that can negatively affect the recovery process. The elimination of damage as a result of tested immediate response action in a crisis situation that disrupts core IT operations constitutes the aimed advantage of the proposed contribution.

Keywords: *Business Continuity Management, Business Function Criticality, Use Case Points, Business Continuity Testing Points.*

1. Introduction

The present enterprise operations are fully dependent on information technologies, complex software applications and multiple corresponding business functions. Thus, one of the most important, and nowadays obligatory, tasks of the modern enterprises and organizations is the development and the establishment of an efficient and effective *Business Continuity Management* [1]. It is by many experts considered to be one of the key Areas of ICT Competencies [2]. Its imperative establishment in terms of enterprise operational policy and strategy stems from multiple and various unexpected and forecasted physical, human or even technical threats that many organizations have experienced within the recent years.

Immediate system recovery for minimization of operational and financial damage stems from a cautiously documented Business Continuity Plan (BCP) [3] [4]. The BCP specifies the methodologies and procedures required to backup and recover every functional units of the business [5]. Thus

regular testing in a simulative but as real as possible mode of the most critical business functions should be scheduled for controlling the stable functionality of IT business functions in crisis situations. Criticality estimation of business functions and corresponding applications is a vital task to be solved by IT managers when planning business continuity testing exercises that concern their unit. The organization and implementation of successful business recovery tests presuppose the creation of a detailed and accurate documentation of the critical business functions and corresponding applications to the *Business Impact Analysis* (BIA) template [3] [4].

Primary discussed and analyzed topics of the present work are all placed within the context of *testing important IT business functions' rapid recovery according to a hypothetical simple, average or complex testing crisis scenario decided on the basis of corresponding process criticality level.*

The author's attempt is to develop and propose a standard methodology for defining the most crucial IT business functions and processes in order to thoroughly document them in the Business Impact Analysis document and subsequently propose the appropriate recovery test type after calculating and defining exactly their criticality level. The determination of this level is based on consideration of corresponding impact value level [6] and additionally the minimum of business functions which must be up and running to ensure the company's basic operation in emergency events. The applied theory behind the construction of the new approach is the Use Case Points method [7]. Use Case model is one of the most tested in practice methodologies for defining user requirements [8]. Additionally, the Business Continuity Management document core characteristic is that it constitutes a requirements' document as well, though from the scope of business function testing procedure within a crisis situation. Required task by the organizations is the recovery of a business function within the desired Recovery Time Objective (RTO) or even the Maximum Acceptable Outage

[4] [3], as they are determined by the Business Continuity team and documented in the BIA template.

Thus, by implementing the Use Case concept can help IT managers perform accurate and objective business function criticality scoring and respectively determine the testing approach of each function, according to the BIA process documented data. The scoring of function criticality is applied according to the rules of the Use Case Points methodology by considering multiple factors, both technical and environmental which ultimately affect the business function recovery process and the timeframe that it can be accomplished. The estimated effort required to “bring back to life” a vital IT business function will comprise of the key indicator of the corresponding testing approach intended to be applied to the specific function.

2. Testing critical IT business functions within a defined enterprise Business Continuity Management policy

From the point of view of scientific literature and industrial practice BC addresses questions of how to handle risk issues in the case that critical business processes of an organization fails [9]. Historically, BC addressed IT processes, later on, business processes came up as the final purpose of their supporting IT processes [10] [1].

The importance of Business Continuity Testing is outlined and thoroughly analyzed by many experts. Precisely it is stated that organizing regular exercises such as desktop and simulation drills is the only way to discover gaps and address them [11].

The current paper focuses on the IT department’s successful documentation and testing of the most critical functions and processes; hardware and software should support critical business functions, so the IT functions, in large part, will be driven by all the other departments. HR might say “we have to have our payroll application”; marketing might say “without our CRM system, we can’t sell any products”; manufacturing might say “without our automated inventory management system, we can’t even begin to make anything.” Therefore, the IT department’s critical business functions are driven externally, to a large degree [12]. However, the successful IT business continuity management policy should focus to the immediate recovery of the indeed most important operations of the enterprise, defined by the ISO 22301:2012 as Minimum Business Continuity Objective, briefly stated as MBCO [4].

Creating a continuity plan is a long-term process and companies should review the existing documentation as an ongoing project [13] [14]. The actual purpose of testing is

to achieve organizational acceptance that the solution satisfies the recovery requirements. Plans may fail to meet expectations due to insufficient or inaccurate recovery requirements, solution design flaws or solution implementation errors [3]. The differentiation of critical (urgent) and non-critical (non-urgent) organization functions/activities is the core task of B.I.A. Critical functions are those whose disruption is regarded as unacceptable. Acceptability is mainly affected by the cost of recovery solutions. A function may also be considered critical in the case that it is imperative due to specific law.

2.1 Exercise Categories

The exercises as they were defined by the Business Standards Institute are divided into 3 basic categories [3]:

Tabletop Exercises: they typically involve a small number of people and participants, who work through a simple scenario, discuss specific aspects of the plan and only a few hours are consumed.

Medium Exercises: conducted within a "Virtual World" and bring together several departments, teams or disciplines.

Complex Exercises: also occur within a virtual world, but maximum realism is essential and duration is unknown.

The results of insufficient and poor testing of software applications are known and obvious within the enterprise environment. Test engineering is seldom planned for in most organizations and as a result, products enter the market insufficiently tested. Negative customer reactions and damage to the corporate image is the natural consequence [15]. Similarly, the test engineering process for critical business functions from the business continuity management standpoint is essential, since the negative enterprise effects caused by unsuccessful response to a real crisis event will be an established fact. Consequently, according to the above statement all business functions should be tested regularly so that all involved staff will be prepared for a real crisis event. The idea behind the proposed contribution is that test success is based on mapping of each IT business function to the most suitable of the above analyzed exercise levels after determination of its impact value level. The way that the mapping is performed is depicted at the following section (see Table8).

2.2 Impact Value Levels of IT business functions

Darril Gibson [6] indicates the impact value level of each business function according to its accepted downtime period without causing negative effects to the enterprise or the organization. The four levels of impact value are:

Level 1: business functions must be available during all business hours. Online systems must be available 24 hours per day and 7 days per week.

Maximum Acceptable Outage (MAO) = 2 hours
 Recovery Time Objective (RTO) < 2 hours

Level 2: business processes can survive without the business function for a short amount of time.

Maximum Acceptable Outage (MAO) = 24 hours (1 day)
 Recovery Time Objective (RTO) < 24 hours

Level 3: business processes can survive without the business function for one or more days.

Maximum Acceptable Outage (MAO) = 72 hours (3 days)
 Recovery Time Objective (RTO) < 72 hours

Level 4: business processes can survive without the business function for extended periods.

Maximum Acceptable Outage (MAO) = 168 hours (1 week)
 Recovery Time Objective (RTO) < 168 hours

The above mentioned levels will be applied to the new proposed Business Continuity Testing Points methodology so that the responsible IT manager of a specific business function will be able to classify it to the appropriate exercise category.

3. The Use Case Points

Use Case points method [16] [17] [18] [15] is especially valuable in the context of early size measurement and effort estimation, because it employs use cases as an input [19], Use cases, proposed by Jacobson [20] [21], are a popular form of representing functional requirements. Moreover, according to the survey conducted by Neill and Laplante in 2003 [22], 50% of projects have their functional requirements presented as scenarios or use cases. They are also available in the early stages of soft-ware development. Thus, due to above stated important feature of functional requirements documentation and taking into consideration that the Business Continuity strategy is based on business functions recovery action requirements, Use Case method can comprise of a new approach to defining exact documentation in Business Impact Analysis documents and furthermore an assisting tool of estimating precisely the Recovery Time Objective and Maximum Acceptable Outage. As it was already mentioned the reason and the need for introducing the new method is to avoid the manual vague estimation of these values based solely on IT manager's practical experience. Before analyzing the new proposed model of Business Continuity Testing Points, a brief reference to the Use Case Points method is required.

3.1 Classifying Actors and Use Cases

The primary step of the Use Case Points procedure is to classify and calculate the Actors' weights and the weight of a Use Case. Classification method with regard to complexity degree of both Actors (see Table1) and Use Cases (see Table2) are listed respectively.

Table 1: Actor's Classification

<i>Actor Type</i>	<i>Weighting Factor</i>
Simple	1
Average	2
Complex	3

Simple Actor [7] [23] [24] represents another system with a defined Application Programming Interface, API, an average actor is another system interacting through a protocol such as TCP/IP, and a complex actor may be a person interacting through a GUI or a Web page.

By counting the number of actors of each kind (complex, average or simple), multiplying each total by its weighting factor and finally adding up the products we calculate the total Unadjusted Actor Weights, briefly mentioned as UAW. The result of the calculation is provided by Eq. (1):

$$UAW = \sum_{i=1}^n A_i * W_i \quad (1)$$

where n= Number of Actors, A= Actor, W= Actor's Weighting Factor.

In a similar way the Unadjusted Use Case Weight (UUCW) value is calculated by multiplying number of each use case category by the corresponding weighting factor, and the products are added up according to Eq. (2),

$$UUCW = \sum_{i=1}^n U_i * W_i \quad (2)$$

where n= Number of Use Cases, U= Given Use Case, W= Use Cases weighting factor

The UAW is added to the UUCW to get the unadjusted use case points from Eq. (3)

$$UUCP = UAW + UUCW \quad (3)$$

Table 2: Use Case Classification

<i>Use Case Type</i>	<i>No of Transactions</i>	<i>Weighting Factor</i>
Simple	≤3	1
Average	4 – 7	2
Complex	>7	3

3.2 Evaluation of Technical and Environmental Factors

In Use Case Points methodology apart from the computation of UUCP value, various Technical (see Table3) and Environmental Factors (see Table4) are considered and computed with respect to Software Application complexity.

After their computation the Adjusted Use Case Points (UPC) are calculated with the help of a special equation, in which Unadjusted Use Case Points value, Technical Complexity Factors (TCP) value and Environmental Factors (EF) value are multiplied.

Table 3: Technical Complexity Factors in Use Case Points

<i>Factor</i>	<i>Description</i>	<i>Weight</i>
T1	Distributed System	2
T2	Response adjectives	2
T3	End – User efficiency	1
T4	Complex Processing	1
T5	Reusable Code	1
T6	Easy to install	0.5
T7	Easy to Use	0.5
T8	Portable	2
T9	Easy to change	1
T10	Concurrent	1
T11	Security features	1
T12	Access for third parties	1
T13	Special Training Required	1

Table 4: Environmental Factors in Use Case Points

<i>Factor</i>	<i>Description</i>	<i>Weight</i>
F1	Familiar with RUP	1.5
F2	Application Experience	0.5
F3	Object – Oriented experience	1
F4	Lead Analyst capability	0.5
F5	Motivation	1
F6	Stable requirements	2
F7	Part – time workers	-1
F8	Difficult programming language	2

The formula applied for calculating Technical Complexity Factor (TCF), is provided by Eq (4):

$$TCF = 0.6 + (0.001 * TFactor) \quad (4)$$

after multiplying the value of each Technical Factor (see Table3) by its corresponding weight and then adding all these numbers to get the sum called TFactor.

In the same way Eq. (5) is applied for calculating Environmental Factor (EF):

$$EF = 1.4 + (-0.03 * EFACTOR) \quad (5)$$

after multiplying the value of each Environmental Factor (see Table4) by its corresponding weight and then adding all these numbers to get the sum called EFactor.

The final calculation of the Adjusted Use Case Points (UPC) is provided by Eq. (6):

$$UPC = UUCP * TFC * EF \quad (6)$$

The estimation effort is the final part of the Use Case Points method. By multiplying the specific value (man-hours) by the UCP, estimated effort can be obtained [23] [25]. A factor of 20 man-hours per UCP for a project is suggested by Karner [7].

4. The Business Continuity Testing Points Method (BCTP)

4.1 Motivation, utilized theory and algorithmic process behind the BCTP methodology

The basic motivation for the construction of the BCTP approach as to defining exact exercise category for each IT function and enabling the implementation of effective recovery tests according to the defined RTO and MAO timeframes of business impact analysis documentation, has been the elimination of erroneous BIA documentation of high priority functions and processes which usually leads to poor testing implementation. As it was already mentioned, the basic reason is that RTO and MAO timeframes documentation in a BIA template is based on employers' everyday operational experience within the organization.

The following step was the selection of the appropriate existing and proved by practical implementation theory in order to derive the new model. With the above stated assumption, since the BIA documentation constitutes a basic element of the general Business Continuity Plan, which is considered to be a requirements' document from system recovery standpoint, a standard requirement analysis and practically implemented methodology such as Use Case Points approach was required as a driving method for the construction of the Business Continuity Testing Points approach. Moreover the testing goal is to satisfy organizational acceptance that the solution meets the recovery requirements.

As in the case of the Use Case Points methodology, the proposed model's construction is separated in 2 parts. The first part constitutes the mapping of Actors and Use Cases for calculation of Unadjusted Use Case Points, to the Actor Types 1 (Human Level) and 2 (Application Level) and Business Processes, instead of Use Cases, in order to calculate Unadjusted Business Function Testing Points (UBFTP). The idea is that in case of obvious and simple business function with low score of unadjusted points there will be no need for further analysis, and thus direct decision about impact value level and exercise category will be made in the final 2nd part. It is assumed that this simple approach concerns only business functions which are not included in MBCO and the corresponding exercise category planned will be either tabletop or medium. Exercise category is based on the derived number of unadjusted points and impact value level (3 or 4).

The second part constitutes a process intended for more complicated IT business functions for which complex exercises should be planned. Thus, exact impact value level should be decided first. The idea is the utilization of technical recovery, environmental recovery and also

unexpected recovery factors which will enable better understanding of the exact impact value level (1 or 2) of the business function. The term *Unexpected Recovery Factors* is a new additional factor category included in the proposed model since Business Continuity Tests should be characterized by hypothetical unexpected conditions that will could prolong the recovery process.

Consequently the algorithmic steps to be implemented in terms of classifying IT business functions to the appropriate exercise category defined by the Business Standard Institution are the following:

Step 1: Defining Actor Types of both levels (Human and Application)

Step 2: Counting Unadjusted Actor Weights of Type 1, which are named as Unadjusted Human Weights, briefly mentioned UHW and Unadjusted Actor Weights of Type 2 the so called Unadjusted Application Weights, briefly stated UAPW. The Total number of Unadjusted Actor Weights is (TUAW) is provided by adding up the weight values of the two Actors.

Step3: Compute Unadjusted Business Process Weights (UBPW)

Step 4: Compute Unadjusted Business Function Recovery Points (UBFRP)

Step 5: Define Impact Value Level and determine whether business function is included in the MBCO, by considering value of UBFRP

Step 6: If Function is not included in MBCO then Impact Value Level is 3 or 4 and Exercise category is *tabletop* or *medium*. Exact definition of levels and exercise categories is not important since it is considered that enterprise can survive without the specific function for a few days. However, if exact definition of the above elements is desired by the organization, the process is the same as it is in the case of complex IT functions that are included in MBCO.

Step7: If Function is included in MBCO then exact impact value level (1 or 2) must be defined. Determined exercise category is complex. The exact Impact Value level is calculated by considering Technical Recovery Factors, Environmental Recovery Factors and Unexpected Event Factors. Impact value level depends on the Adjusted Business Function Recovery Points (ABFRP) value and the total Recovery Effort value that will be computed. For better understanding of the analyzed algorithmic steps, a schematic presentation is also included in the current work (see Fig.1) via a UML Activity Diagram.

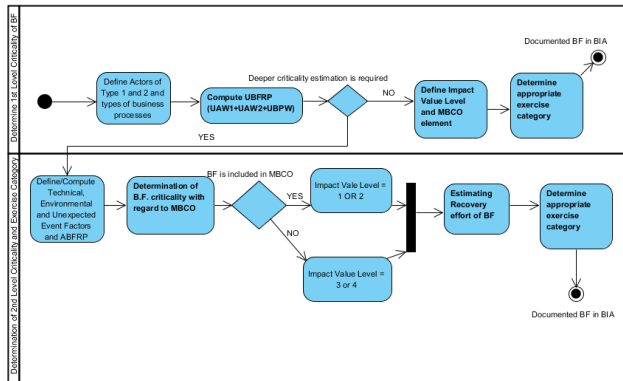


Fig. 1 Activity Diagram of the proposed BCTP approach

4.2 Definition of Business Continuity Testing Points Actor Types, Unadjusted Points and Factors

The first part of the model includes the defined mapping of Use Case Points parameters to the corresponding Business Continuity Testing Points parameters and also the calculation of the Unadjusted Points. The Actor classification of the former case is replaced by 2 different Actor Type classification in the new model. Actor Type 1 represents Human participation in the business function or process (see Table5). Moreover, Actor Type 2 represents the application involved in the same function or process (see Table6).

Table 5: Actor Type 1 BCTP Classification

<i>Actor Type 1 (Human Level)</i>	<i>Weighting Factor</i>
Basic responsibility tasks	1
Average responsibility tasks	2
Complex responsibility tasks	3

Table 6: Actor Type 2 BCTP Classification

<i>Actor Type 2 (Application Level)</i>	<i>Weighting Factor</i>
Simple Tasks	1
Average	2
Complex	3

Eq. (7) and Eq. (8) respectively provide the results of the calculation of each Actor type

$$UHW = \sum_{i=1}^n A1_i * W_i \quad (7)$$

$$UAPW = \sum_{i=1}^n A2_i * W_i \quad (8)$$

The Total Unadjusted Actor Weights (TUAW) is calculated by Eq. (9)

$$TUAW = UHW + UAPW \quad (9)$$

Additionally, Use Cases are now replaced by *Business Processes*. We do not count Use Case Unadjusted weights but Business Process Unadjusted Weights. The business process complexity is now classified according to the number of process steps (see Table7). The calculation of the Unadjusted Business Process Weights is implemented with Eq. (10)

$$UBPW = \sum_{i=1}^n BP_i * W_i \quad (10)$$

Table 7: Business Process Classification in BCTP

<i>Business Process Type</i>	<i>Number of Process Steps</i>	<i>Weighting Factor</i>
Simple	<=3	1
Average	4 – 7	2
Complex	>7	3

Finally through Eq. (11) the Unadjusted Business Function Recovery Points (UBFRP) value can be calculated,

$$UBFRP = TUAW + UBPW \quad (11)$$

By obtaining the value of unadjusted points the first level evaluation of function criticality has been terminated. Functions that are not urgent can be simply documented to BIA template either with Impact Value Level3 and Medium Exercise Plan or with Impact Value Level4 and Tabletop Exercise plan (see Table8).

Table 8: IT Business Function Impact Value Levels and corresponding planned Exercise Category

<i>Business Function</i>	<i>Impact Value Level</i>	<i>Exercise Category</i>	<i>Included in MBCO (urgent)</i>
BF1	Level 1	Complex	YES
BF2	Level 2	Complex	YES
BF3	Level 3	Medium Exercise	NO
BF4	Level 4	Tabletop Exercise	NO

However, for urgent functions the above mapping is proposed so that disaster recovery testers will focus and test thoroughly whether the recovery time spent during the exercise meets the Rational Time Objective (RTO) of the function or its relative Maximum Acceptable Outage (MAO). In the case that none of the above hypothesis will be fulfilled the aforementioned values of the specific function should be reconsidered through a reengineering procedure. To calculate the effort spent for testing recovery plan and compare it to the RTO and MAO expected values the Technical, Environmental and Unexpected factors should be calculated as well throughout the second level of criticality estimation procedure.

4.2 2nd Level Business Function Criticality Estimation - Mapping Use Case Points Factors to BCTP Factors for detailed evaluation of business function

The model's second part, is currently under development. Throughout this part a deeper 2nd level estimation of function criticality is aimed via calculation of crucial factors. Environmental and Unexpected Factors, and additionally the corresponding weight of each factor are subject of future research. However the above mentioned factors have been already derived, but modified for the purposes of the new approach, from the Use Case Points model and they are listed in the following subsection (see Table9, Table10)

Table 9: Technical Recovery Factors in BCTP method

<i>Factor</i>	<i>Description</i>
TRF1	Application's communication with other systems
TRF2	Function Type
TRF3	User's skills
TRF4	Complex functions
TRF5	Routine functions
TRF6	Easy to restore
TRF7	Easy to process
TRF8	Installed locally or in remote server
TRF9	Exists alternative application (i.e. older)
TRF10	Functional Area
TRF11	Security features
TRF12	Utilized by third users
TRF13	Extreme and special knowledge required

Table 10: Environmental Recovery Factors in BCTP Method

<i>Factor</i>	<i>Description</i>
ERF1	Familiar with Business Recovery procedures
ERF2	Users' application experience
ERF3	Users' recovery task knowledge
ERF4	Leader's capability
ERF5	Team's motivation
ERF6	Stable requirements of system's recovery level (Stable MBCO)
ERF7	Part – time personnel
ERF8	Customers' needs direct effect

Apart from the Technical and Environmental factors which are derived according to the Use Case Points methodology, the concept of *Unexpected Recovery Factor(s)* is introduced, enriching the state of the art in Business Continuity of IT functions (see Table11). The consideration and calculation of the specific factors is considered to be highly important; the negative influence of such factors could result to time deviation from the desired RTO and MAO of a business function recovery effort. The equation that will provide the URFactor, and the weight of each individual factor constitute subject of further research. However a short reference to each unexpected factor is required.

Weather conditions can constitute preventive element from reaching the recovery site or special office. The *Disaster Type* should also be considered in a testing scenario. Slight system interruption differs from building collapse in terms of recovery time. Additionally *Timely Information Distribution of Crisis Event* can result either to immediate response of staff or late response if the message is sent later on.

Table 11: Unexpected Recovery Factors in BCTP Method

<i>Factor</i>	<i>Description</i>
URF1	Weather conditions
URF2	Disaster Type
URF3	Timely Information Distribution of Crisis event
URF4	Urban conditions
URF5	Staff availability
URF6	Network availability

Traffic, closed roads and similar *urban conditions* can also badly affect the recovery as in the case of bad weather conditions. Moreover, *staff availability* in case of sickness and *network availability* due to technical reasons can trigger important deviation from the expected recovery time.

The final step of the BCTP model includes the calculation of the Adjusted Business Function Recovery Points (ABFRP). The value will be provided by multiplication of Unadjusted Points value, Technical Recovery Factors, Environmental Recovery Factors and Unexpected Recovery Factors according to the Eq. (12).

$$ABFRP = UBFRP * TRF * ERF * URF \quad (12)$$

The above number will be considered towards the calculation of the Recovery Testing Effort (RTE) of a unique IT business function. The value of the effort should be less than or equal to the desired RTO, or by the worst case scenario equal to MAO. In any other case recovery tests or RTO, MAO values must be modified. The equation and method for calculating Recovery Time Effort is also a future feature of the model.

4. Conclusions

IT Business continuity constitutes a crucial issue of modern enterprises. The tests performed in terms of preparedness against crisis situations aim to ensure the immediate and almost continuous operation of minimum demanded IT business functions. The current work includes the proposal of a new method of planning efficient and effective recovery tests derived from the Use Case Points concept which is entitled as Business Continuity Testing Points. The accuracy of the new model is expected due to calculation of various factors that can trigger time deviation of system's recovery from the expected Recovery Time Objective or the Maximum Acceptable Outage. The current paper analyzes the initial version of the contribution. So far extension of the Use Case Points model by modifying factors, creating new actor types and introducing the Unexpected Recovery Factors has been the model's primary task. Future work and research will focus on definition of Weights of the Technical, Environmental and Unexpected Recovery Factors, development of standard mathematical equations for counting the above stated factors and creation of the final equation which will calculate the recovery effort for an individual IT business function in order to compare it with desired RTO and MAO values and finally schedule the most appropriate recovery exercise category.

Acknowledgments

This work was supported by ESF operational program "Education for Competitiveness" in the Czech Republic in the framework of project „Support of engineering of excellent research and development teams at the Technical University of Liberec“ No. CZ.1.07/2.3.00/30.0065.

References

- [1] Y. Asnar and P. Giorgini, "Analyzing business continuity through a multi-layers model," in BPM '08: Proc. of the Int. Conf. on Business Process Management, 212-227. Springer, 2008.
- [2] K. Antlova, L. Popelinsky and J. Tandler, "Long term growth of SME from the view of ICT competencies and web presentations," *Ekonomie a Management*, pp. 125-139, 2011.
- [3] Business Standard Institute, *Exercises for excellence*, 2008.
- [4] Business Standard Institute, *BS ISO 22301 : 2012*, 2012.
- [5] Doughty, K., ed., *Business Continuity Planning Protecting Your Organization's Life.*, Best practice series. Auerbach, Boca Raton, 2001.
- [6] D. Gibson, *Managing Risks in Information Systems*, Jones & Bartlett Learning, 2010.
- [7] G. Karner, "Resource Estimation for Objectory Projects," in *Systems SF AB*, 1993.
- [8] A. Cockburn, *Writing Effective Use Cases*, Boston: Addison- Wesley Longman Publishing Co. Inc., 2000.
- [9] W. Cassel, "Assessing business continuity requirements," in *Power Engineering Society General Meeting*, 1: 822-823. IEEE Computer SOCIETY, 2005.
- [10] E. Zambon, D. Bolzoni, S. Etalle and M. Salvato, "A model supporting business continuity auditing and planning in information systems," in *ICIMP '07: Proceedings of the Second International Conference on Internet Monitoring and Protection*, 33. IEEE Computer Society, 2007.
- [11] C. Sapriel, "Effective crisis management: Tools and best practice for the new millennium," *Journal of Communication Management*, Henry Stewart Publications 1, 2003.
- [12] S. Snedaker, *Business Continuity and Disaster Recovery Planning for IT Professionals*, Burlington: Syngress Publishing Inc., Elsevier Inc., 2007.

- [13] R. J. Kepenach, "Business Continuity Plan design," in ICIMP '07: Proceedings of the Second International Conference on Internet Monitoring and Protection, 27. IEEE Computer Society, 2007.
- [14] W. Lam, "Ensuring Business Continuity," in IT Professional, 4: 19-25, 2002.
- [15] S. Nageswaran, "Test Effort Estimation Using Use Case Points," in Quality Week 2001, San Francisco, California, USA, 2001.
- [16] C. Nagar and A. Dixit, "Efforts Estimation by Combining the Use Case Point and COCOMO," International Journal of Computer Applications, vol. 52, no. 7, 2012.
- [17] S. Diev, "Use cases modeling and software estimation: Applying Use Case Points," ACM SIGSOFT Software Engineering Notes, vol. 31, no. 6, 2006.
- [18] M. W. Kamal, M. A. Ahmed and M. El-Attar, "Use Case-Based Effort Estimation Approaches: A Comparison Criteria," in Software Engineering and Computer Systems - Second International Conference, ICSECS 2011, Proceedings, Part III , Kuantan, Pahang, Malaysia, 2011.
- [19] M. Ochodek, J. Nawrocki and K. Kwarciak, "Simplifying effort estimation based on Use Case Points," Journal of Information and Software Technology, Elsevier, vol. 53, no. 3, pp. 200-213, 2011.
- [20] I. Jacobson, Object-oriented development in an industrial environment, ACM SIGPLAN Notices 22 (122), 1987.
- [21] I. Jacobson, P. Cristerson, P. Jonsson and G. Overgaard, Object-Oriented Software Engineering: A Use Case Driven Approach, Addison-Wesley Longman, Inc., 1992.
- [22] C. Neill and P. Laplante, "Requirements engineering: the state of the practice," Software, vol. 20, no. 6, pp. 40-45, 2003.
- [23] G. Banerjee, "Use Case Points -An Estimation Approach," 2001.
- [24] S. Kusumoto, F. Matukawa and K. Inoue, "Estimating Effort by Use Case Points: Method, Tool and Case Study," in Proceedings of 10th International Symposium on Software Metrics (METRICS'04), 2004.
- [25] G. Schneider and J. Winters , Applying Use Cases: A Practical Guide, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1998.
- [26] D. Elliot, E. Swartz and B. Herbane, "Just waiting for the next big bang: business continuity planning in the UK finance sector.," Journal of Applied Management Studies., vol. 8, pp. 43-60, 1999.
- [27] C. Brandt, F. Hermann and J. F. Groote, "Generation and evaluation of business continuity process using algebraic graph transformation and the mCRL2 Process Algebra," Journal of Research and Practice in Information Technology, vol. 43, no. 1, 2011.
- [28] M. Blyth, "Business Continuity Management: Building an Effective Incident Management Plan," John Wiley & Sons, 2009.

Athanasios Podaras received a Ph.D degree (2010) in Information Management, an MSc Diploma (2005) in Informatics from Czech University of Life Sciences in Prague, Department of Information Engineering (Czech Republic), and a Bachelor Degree in Agriculture from Technical University of Kalamata – Greece (2001). He is currently a Post - Doctoral researcher in the Technical University of Liberec, Informatics Department. His research interests include areas such as Business Process Requirement Analysis, Business Continuity Management and IT System Recovery in crisis situations, Development of Early warning Information Systems, Applied Information Technologies in Agriculture.

Tomas Žižka has been working since 2006 as a lecturer at the Economics Faculty of the Technical University in Liberec in the Department of Computer Science. He is currently studying a PhD study program System Engineering and Informatics in field of study called economic informatics. Title of the dissertation is the possibility of targeted receiving in non-standard situations. His research interests include areas such as Information Management in the Event of Traffic Accidents, Distribution of Information in Case of Non-standard Situations, Development of Early warning Information Systems

A Metric Based Approach to Extract, Store and Deploy Software Reusable Components Effectively

Muhammad Ilyas¹, Mubashir Abbas², Khansa Saleem³

¹Department of Computer Science and Information Technology, University of Sargodha
Sargodha, Punjab, Pakistan

²Department of Computer Science and Information Technology, University of Sargodha
Sargodha, Punjab, Pakistan

³Department of Computer Science and Information Technology, University of Sargodha
Sargodha, Punjab, Pakistan

Abstract

Software reusability is a valuable methodology for quality, economical and timely software development. The effective use of reusability gives benefits in the form of less time, efforts and cost for quality software development. Reusability also helps to diminish the risk associated with software development and success. Due to inevitable payback, reusability has grown up to be most accepted practice for software development. But reusability handling methods and techniques are not well organized, so need is to formalize the reusability process in order to get its actual benefits in form of time, cost and effort savings. Formal and structured approach is required in reusability practices because reusability observation, extraction, classification and deployment methods are not disciplined. In this paper a framework is suggested to make reusability process formal and organized. In this approach quality earning criteria are defined at each level of reusability process to observe the need of reusability, extracting reusable components, classifying them and then integrating with new systems efficiently. The aim is to formalize each activity of reusability process to get satisfactory and quality results.

KEYWORDS: *Software Component Assessment, Software Reusability, Software Risk, Reusability Process.*

1. Introduction

Reusability methodology has turn into a productive tool for software development as it reduces time, cost and work required for the software development [1, 2, 3]. Reusability increases reliability, quality and productivity of software products by using already existing tested components [1, 6, 5]. Developers consider it most favorable choice for economical development of business and technical projects [1]. Reusability ensures within time delivery of software products and minimizes risk involved with its success [2, 4]. Reusability boosts up confidence of developing team as they have previously groundwork in the current domain [2]. Reusability gives benefits in the form of portability, maintainability and productivity improvements [1, 4]. Due to inevitable paybacks reusability is widely used for cheaper and

timely software development purposes. But techniques for its observability, suitability, selection and deployments are not well organized, so need is to apply structured and formal approach in the implementation of reusability to acquire real benefits.

This study aims to formalize the all activities performed during reusability life cycle [1]. It will help to obtain satisfactory and quality results. For this purpose the reusability life cycle is divided into three stages: Reusable Component's Extraction, Reusable Component's Storage and Reusable Component's Deployment. Certain metrics are defined at each stage aiming quality and productive output from each stage. Meticulous meditation is given to the extraction of quality software components that can be reused for productive software development purpose. Than these reusable components are stored and classified in the reuse repository on the basis of certain characteristics. These characteristics make searching and retrieval process efficient [7]. At the end Pre Adopt test is conducted to find the most suitable and trusted reusable components according to the new system requirements [3].

The next section constitutes related work followed by a proposed process and conclusion and future work.

2. Related Work

Al-badareen [1] proposed framework consists of extraction, adaptation, storage, pre-store and pre-use process. The extraction of reusable components is performed during develop for-use process, which is focused on that how to extract suitable information for a reusable component. The storage process is performed to store extracted reusable components in the library, so they can be accessed easily for reusable purposes. The adaptation process is how to pick a suitable component according to new system requirements. Pre-store is the process to evaluate and enhance reusable components according to certain standards to satisfy library requirements. Pre-store

process evaluates reusable components on the basis of co-existence, compliance, generality and adaptability characteristics. The pre-use process is the process to evaluate and modify the reusable components, retrieved from the library to satisfy the new system requirements. Three main characteristics: suitability, accuracy and compliance are considered and evaluated in pre-use process. Both pre-store and pre-use processes include a reusability test. Reusability test checks for certain library conditions to be fulfilled by the reusable component during pre-store process. In pre-use process, reusability test evaluates the reusable components according to certain requirements of new system.

G. Sindre [7] has introduced reboot approach to deal with reusability issues. Reboot qualification model is used to ensure existence of quality reusable components in the library. Portability, flexibility, understandability and confidence metrics are used for evaluation of reusable components. It ensures that quality components have qualified for storage in the library. With large libraries having large numbers of reusable components of different domains, a proper way is required for efficient retrieval of these reusable components. For this purpose reboot used a faceted classification scheme to classify these reusable components in the library. He made use of four facets abstraction, operations, operates on and dependencies for the classification of reusable components. Also some other attributes like who developed it, when it was developed and how big it is, are considered for storing and classifying the reusable components in the library.

Fazl-e-amin [6] developed the reusability attribute model and metrics to measure the attributes of reusable components. This model defines those attributes of an aspect oriented component that contributes to its reusability. Flexibility, maintainability, portability, scope coverage, understandability and variability were identified as the attributes of the given model.

3. Proposed Framework

Although Reusability has grown up into a fruitful practice for software development but its implementation techniques are not systematic and formal. Formal approach is pretty required in reusability observation, extraction, classification and deployment processes. To overcome limitations in usage of reusability, a frame work is purposed in this paper. Aim is to formalize the reusability process at each step that will help to gain quality results from each stage of reusability life cycle [1].

According to this purposed approach reusability life cycle is divided in to three stages Reusable Component's Extraction, Reusable Component's Storage and Reusable Component's Deployment shown in fig. 1.

At each stage certain metrics are defined, which component has to be fulfilled to move on to the next stage during reusability life cycle. At Extraction stage components that can be reused are extracted. These reusable components must be capable of some particular characteristics to become an effective reusable component [1]. When a reusable component is selected then it needs to be store into the reusable repository from where it can be retrieved for reuse purpose [1, 7]. So searching and retrieval of reusable components can be easily made if components are properly stored. Third stage of this framework is deployment process in which a component of similar functionality is fetched from the Reuse Repository. Now this fetched component is assessed for suitability with new system's requirements. If it satisfies the new system requirements then it is adopted and put for Pre Adopt test for final selection [1]. After using metrics at each stage quality results are anticipated aiming to get true benefits of reusability in the form of less time, cost and work for the quality and productive software development.

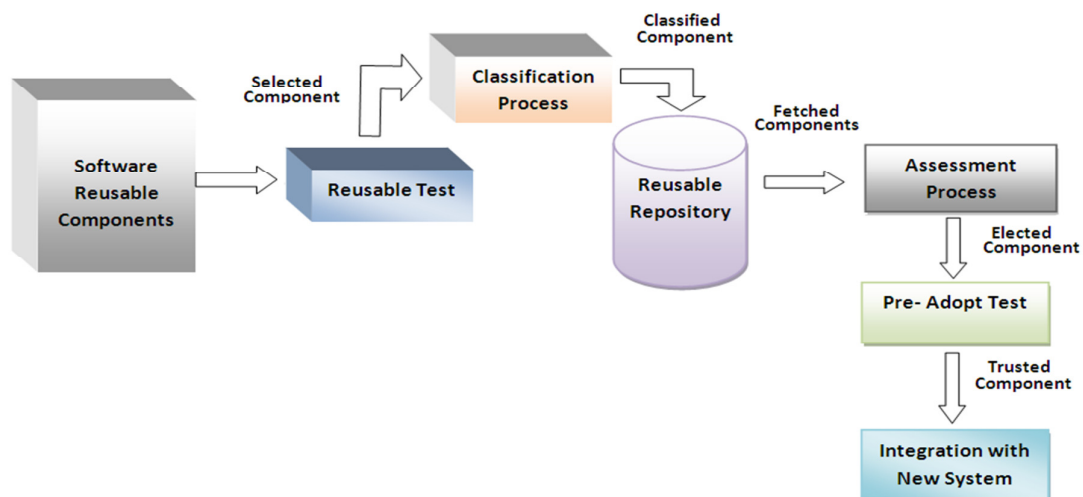


Fig.1 RESAD Framework

4. Extraction of Reusable Components

Reusability has become an ideal methodology for software development as it helps to develop quality software components with minimum cost and effort. It is possible when there is some repository of reliable, portable and qualitative components that can be reused into a new system. In this way time, cost and effort is decreasing and productivity of the new system is increasing [1, 3, 5, and 6]. So there must be a formal process to extract the components that can be reused. Extraction of reusable components is the process of finding components that can be reused. This selection is made on the basis of some particular characteristics shown in fig.2. These characteristics are defined to produce healthy reusable components. In this study following metrics are defined that a software component must fulfill in order to be select as a reusable component.

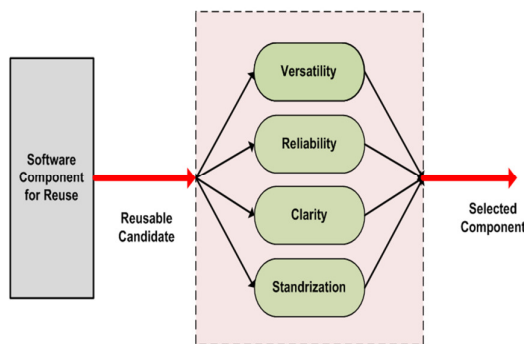


Fig. 2 Extraction Process

4.1. Versatility

Versatility characteristic of software component is the ability to deal with different type of platforms efficiently. It means that entire software component can perform its functionality independent of hardware and software constraints [1, 7]. Component is providing support to wide range of machines on which different operating systems are running.

Versatility of software component also defines that it contains support or compatible with different programming languages. Component with versatility feature is considered more flexible and portable, has the great chance to be reused [7].

Here Versatility Metric is being used as combination of Generality and Portability Metric of software components. To determine versatility of software components both generality and portability metrics will be analyzed. To measure versatility characteristic following metrics are proposed:

4.1.1 Generality Metric

According to W. J. Salamon generality of software can be measured by following metrics [8].

- Multiple Usage Metric

A module is more general if it is referenced by more than one module. Higher the value of entire metric denotes maximum generality of software component.

$$\frac{\text{Total no. of modules referenced by more than one module}}{\text{Total no. of modules}}$$

This relation can be denoted mathematically as follows:

$$M_{US} = \frac{\sum_i^n Ref_i}{n}$$

$$\therefore Ref_i = \begin{cases} 1, & \text{when function is reffernced by more than one module} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2 \dots n$ & $n=$ Total no. of modules (1)

- Mixed Function Metric

A function is considered to be Mixed Function if it performs more than one task like I/O as well as processing operation. Higher value of this metric provides maximum generality.

$$\frac{\text{Total no. of modules with mixed function property}}{\text{Total no. of modules}}$$

Mathematical form of this relation is:

$$M_{FN} = \frac{\sum_i^n Mix_i}{n}$$

$$\therefore Mix_i = \begin{cases} 1, & \text{when function performs more than one task} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2 \dots n$ & $n=$ Total no. of modules (2)

- Data Volume Metric

A module that can process unlimited range of inputs is considered more general.

$$\frac{\text{Total no. of modules with data volume property}}{\text{Total no. of modules}}$$

Mathematically it can be represented as:

$$D_{VM} = \frac{\sum_i^n Vm_Limit_i}{n}$$

$$\therefore Vm_Limit_i = \begin{cases} 1, & \text{when function is data volume limited} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2 \dots n$ & $n=$ Total no. of modules (3)

- Data Value Metric

A module is more general if it can process long range of data items.

$$\frac{\text{Total no. of modules with data value property}}{\text{Total no. of modules}}$$

Mathematical representation of this relation is:

$$D_{VL} = \frac{\sum_i^n Vl_Limit_i}{n}$$

$$\therefore VL_{Limit_i} = \begin{cases} 1, & \text{when function is data value limited} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2 \dots n$ & $n=$ Total no. of modules (4)

- Redefinition of Constants Metric

To change the function of modules constants should not be redefined.

$$\frac{\text{No. of constants that are redefined}}{\text{Total no. of constants}}$$

Mathematical form of this relation is:

$$C_{RF} = \frac{\sum_i^c Rdf_i}{C}$$

$$\therefore Rdf_i = \begin{cases} 1, & \text{when constant is redefined} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, c$ & $c=$ no. of constants (5)

K.K Aggarwal presented the idea of General Programming using templates. According to him templates are more generic and can be used with different data types [9].

Templates can be classified in to two types Function Templates and Class Templates.

- Function Template Metric

To make behaviour of a function general, to be operate on each given data type function templates are used.

$$\frac{\text{No. of fuctions using function template}}{\text{Total no. of functions}}$$

Mathematically it can be represented as:

$$F_{TM} = \frac{\sum_i^n UFt_i}{n}$$

$$\therefore UFt_i = \begin{cases} 1, & \text{when function is using function template} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, n$ & $n=$ Total no. of functions (6)

- Class Template Metric

To make behaviour of a class general, to accept objects of particular data type class template are used.

$$\frac{\text{No. of classes using class template}}{\text{Total no. of classes}}$$

Mathematical form of this relation is:

$$C_{TM} = \frac{\sum_i^c UCt_i}{n}$$

$$\therefore UCt_i = \begin{cases} 1, & \text{when class is using class template} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, C$ & $C=$ no. of Classes (7)

Probable effect of all above described metrics about generality of software component is mentioned in following table.

Table 1: Effect of Generality Metrics

Metric Name	Metric Symbol	Value	Reusability Probability
Multiple Usage	M _{US}	High	High
Mixed Function	M _{FN}	Low	High
Data Volume	D _{VM}	Low	High
Data Value	D _{VL}	Low	High
Redefinition of Constants	C _{RF}	Low	High
Function Template	F _{TM}	High	High
Class Template	C _{TM}	High	High

4.1.2. Portability Metrics

W. J. Salamon made use of following metrics to measure portability characteristic of software components [8].

Software Independence Measurement

- Compatibility Metric

No. of operating systems with software is compatible

$$C_{MP} = \frac{\sum_i^n Os_i}{n}$$

$$\therefore Os_i = \begin{cases} 1, & \text{when operating system is compatible} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, n$ & $n=$ no. of operating systems (8)

- System Utilities Metric

Total no. of system utility utilized. (To measure S/W dependability)

$$S_{UY} = \frac{\sum_{i=0}^n Uty_i}{n}$$

$$\therefore Uty_i = \begin{cases} 1, & \text{when system utility is being utilized} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, n$ & $n=$ no. of operating system utilities (9)

- Standard Constructs Metric

Checks for Common, standard subsets of language used:

$$\frac{\text{Total no. of modules using non standard constructs}}{\text{Total no. of modules}}$$

Mathematical form of this relation is:

$$S_{CN} = \frac{\sum_i^n USing_Nstd_i}{n}$$

$$\therefore Using_Nstd_i = \begin{cases} 1, & \text{when module is using non standard constructs} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, n$ & $n=$ Total no. of modules (10)

Hardware Independence Measurement

- Open system Metric

Are the programming languages and tools (e.g., compilers, DBMS, and user interface) available on other machines? A value of 1 means yes, 0 means No. Mathematically this property can be represented as:

$$O_{SY} = \begin{cases} 1, & \text{when programming languages and tools are available} \\ 0, & \text{otherwise} \end{cases}$$

(11)

- I/O References Metric

Total no. of modules making IO references
Total no. of modules

$$R_{IO} = \frac{\sum_i^n Ref_i}{n}$$

$$\therefore Ref_i = \begin{cases} 1, & \text{when module is making I/O reference} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, n$ & $n=$ Total no. of modules

(12)

- Word/Character Size Metric

Total no. of modules not following convention
Total no. of modules

$$W_{SZ} = \frac{\sum_i^n NT_Conv_i}{n}$$

$$\therefore NT_Conv_i = \begin{cases} 1, & \text{when module is not following convention} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, n$ & $n=$ Total no. of modules

(13)

Anticipated outcome of above metrics measuring Generality of software component is mentioned below:

Table 2: Effect of Portability Metrics

Metric Name	Metric Symbol	Value	Reusability Probability
Compatibility	C _{MP}	High	High
System Utilities	U _{SY}	Low	High
Standard Construct	S _{CN}	Low	High
Open System	O _{SY}	Yes	High
I/O Reference	R _{IO}	Low	High
Word/Character Size	W _{SZ}	High	High

4.2 Clarity

Clarity of the software components is the characteristic which states that component is clear and well understood in its vision, scope and functionality [6, 7]. A component is understandable if it is readable. Then it supports its understandability which enhances analyzability and changeability of software component. Bajeh has suggested following metric to measure readability of software component which is directly connected with its clarity and understandability [10].

- Indentation Metric

No. of lines of code properly indented
No. of lines of code expected to be indented
 Mathematically form of this relation is:

$$N_{LI} = \frac{\sum_i^L Pid_i}{\sum_i^L Eid_i}$$

$$\therefore Pid_i = \begin{cases} 1, & \text{when line of code is properly indented} \\ 0, & \text{otherwise} \end{cases}$$

$$\therefore Eid_i = \begin{cases} 1, & \text{when line of code is expected to be indented} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, L$ & $L=$ Total no. of lines of codes

(14)

- Comments Metric

No. of lines of code commented
Total no. of lines of code

Mathematically it can be represented as:

$$N_{LC} = \frac{\sum_i^L Cmt_i}{L}$$

$$\therefore Cmt_i = \begin{cases} 1, & \text{when line of code is commented} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, L$ & $L=$ Total no. of lines of code

(15)

Anticipated effect of clarity metrics is described in following table.

Table 3: Effect of Clarity Metrics

Metric Name	Metric Symbol	Value	Reusability Probability
Indentation Metric	N _{LI}	High	High
Comments Metric	N _{LC}	High	High

4.3 Reliability

Reliability characteristic of software component states that entire component is a confident component. It has performed its functionality satisfactory without any failure in different environments at different times [1, 7]. Reliability of software components can also be defined as the entire component is capable of excellent quality results in various circumstances without any exception that is associated with quality and performance decreasing.

Mathematically it can be represented as:

No. of time used
No. of time failure is reported

$$R_{LY} = \frac{\sum_i^T Flr_i}{T}$$

$$\therefore Flr_i = \begin{cases} 1, & \text{when failure is reported} \\ 0, & \text{otherwise} \end{cases}$$

Where $i=1, 2, \dots, T$ & $T=$ Total no. of time used

(16)

4.4 Standardization

Standardization characteristic of software component declares that component is developed in view of standard software engineering rules and practices [1]. These rules are necessary to be followed for standard and quality software products developments.

In other words it is the certificate held by the software products that they can be reused for satisfactory results.

Mathematical representation of this metric is as follows:

$$S_{TD} = \begin{cases} 1, & \text{when component is following software standards} \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

5. Storage and Classification of Reusable Components

During extraction process, candidates that are selected as reusable components needs to be store somewhere from which they can be retrieved and accessed easily [1]. So there must be a proper way to store and classify the software components. Now it will be easy to access and retrieve them with minimum effort and time that are the focal benefits of the reusability methodology [7].

In this purposed approach components that are selected as reusable components are classified on the basis of some features shown in fig. 3.

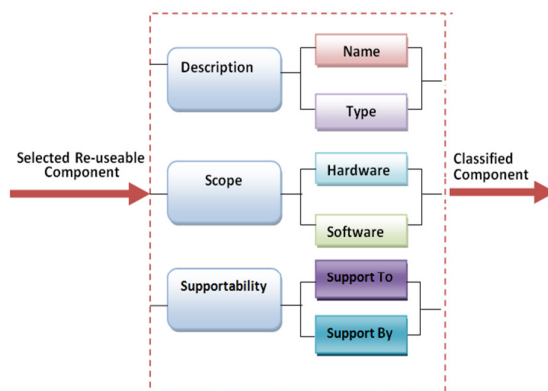


Fig. 3 Storage and Classification Process

5.1 Descriptions

Description attribute of classification process contains information about name and type of reusable component.

- *Name* attribute specifies the name of reusable component.
- *Type* attribute defines whether reusable components is requirement document, design template, source code, test case or else one belongs to SDLC phase[7].

5.2 Scope

Scope attribute of classification process provides details about range of hardware and software.

- *Hardware* attribute specifies the series of machines where this component can be run to perform the prescribed task.
- *Software* attribute specifies the range of operating systems that provide support to entire component to perform its functionality.

5.3 Supportability

Supportability attribute of classification process contains two attributes, Support To and Support By.

- *Support To* attribute defines the software category to which it belongs. Software categories can be system software, application software, embedded software etc.
- *Support By* attribute defines the set of compatible programming languages.

6. Assessment Process

Assessment process is the activity during reusability life cycle in which reusable components are evaluated and analyzed according to the new system requirements [1, 7].

The aim is to find the most suitable component for the given requirement. There is possibility that more than one reusable component can be in the repository for the given requirements. In this situation the task is to select the best component that is most suitable to the given requirements [7].

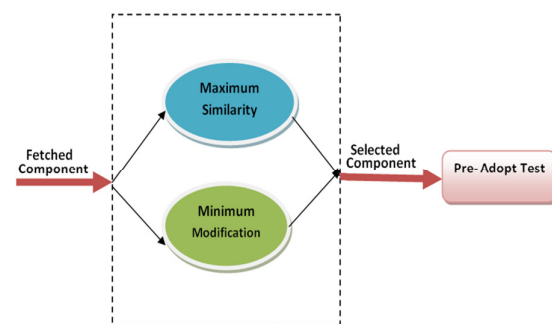


Fig.4 Assessment Process

In this purposed approach this task is performed by Assessment Process. It selects most suitable component from the repository for the given requirements. Assessment is performed on the basis of metrics described in fig. 4:

6.1 Maximum Similarities

The reusable component, having maximum similarity with given requirements is fetched from the repository. The selected component can be the most suitable

choice to carry out the given task efficiently due to the maximum similarities with given requirement. Mathematical form of this relation is as follows:

$$M_{SM} = \frac{\text{No. of features available}}{\text{No. of features required}}$$

$$M_{SM} = \frac{\sum_i^F Ftr_i}{F}$$

$\therefore Ftr_i = \begin{cases} 1, & \text{when required feature is available} \\ 0, & \text{otherwise} \end{cases}$
 Where $i=1, 2, \dots, F$ & $F = \text{Total no. of features}$

(18)

6.2 Minimum Modifications

Assessment Process fetches reusable component from repository which is requiring minimum modifications to meet new system requirements. The selected component with minimum modifications is demanding less efforts, time and cost to perform required functionality.

Minimum Modification metric is compliment of Maximum Similarities metric. Mathematically this relation can be shown as:

$$M_{DN} = \sim M_{SM}$$

(19)

7. Adaptation Process

The assessment process provides components having maximum similarities and minimum modifications to the given requirements. Now this component is most suitable for integrating with new system. But still this component has to cross Pre Adopt test. Pre Adopt test consists of characteristics specifically belongs to new system requirements. These characteristics evaluate that whether the selected component is suitable or not to use in the new system [1].

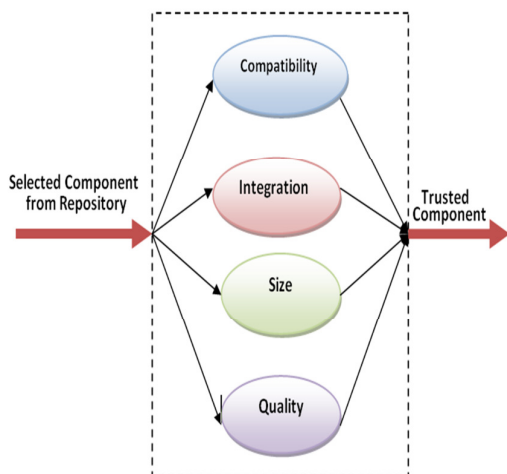


Fig. 5 Pre Adopt Test

7.1 Compatibility

Compatibility metric of Pre Adopt Test evaluates that selected component is compatible with new system. It is an easier job to adjust it into the new system and it can perform required functionality without any complication.

$$C_{MP} = \begin{cases} 1, & \text{when component is compatible with new system} \\ 0, & \text{otherwise} \end{cases}$$

(20)

7.2 Integration

Integration metric of Pre Adopt test analysis that selected component from the reusable repository is easy to integrate with new system. Proper interfaces are available for integrating with other systems. This attribute focuses that overall reusing cost and effort is less than development cost and effort of newly developed component.

$$I_{NT} = \begin{cases} 1, & \text{when interface for integration is available} \\ 0, & \text{otherwise} \end{cases}$$

(21)

7.3 Size

Size metric of Pre Adopt test observes that the size of selected reusable component. It will observe that whole component or some part of it is needed to fulfill given requirement.

$$S_{ZE} = \begin{cases} 1, & \text{when full component is required} \\ 0, & \text{when part of component is required} \end{cases}$$

(22)

7.4 Quality

Quality metric of the Pre Adopt Test determines the quality of selected component. Quality metric depends upon Maximum Similarities Metric, Compatibility Metric and Integration Metric. Positive values of these metrics ensure that selected component is economical and trusted component for subsequent usage. Mathematical relation between Quality and above mentioned metrics can be shown as follows:

$$Q_{TY} = M_{SM} \wedge C_{MP} \wedge I_{NT}$$

(23)

8. Conclusions and Future Work

Reusability methodology has grown up into a productive tool for economical, quality and timely software development. Effective use of reusability improves quality, productivity and maintainability of the software products. But there is need of formal and systematic approach in the use of reusability methodology. This formalism is aiming earning of

actual benefits of reusability in the form of less time, cost and effort.

In this context, a framework is purposed to adopt formalism in the reusable component life cycle. This framework consists of three stages: Extraction, Storage and Deployment of reusable components. Extraction stage extracts the quality based reusable components during software development life cycle. Now this reusable component can be reused for similar kind of problem. Storage Stage classifies and stores these reusable components. After this they can be accessed and retrieved easily. Assessment Process evaluates the reusable component according to new system requirements and finds the most suitable component. In deployment stage Pre Adopt test is conducted for the final selection. After this selected reusable component is integrated with new system. All these tasks are performed on the basis of particular metrics at each level of reusability life cycle.

Our future work is aimed at validation of metrics used in the purposed framework to measure these attributes in different conditions. Prototype of this framework is also under construction, which is the foremost goal of future work.

9. References

- [1] A. B. AL-Badareen, M. H. Selamat and M. A. Jabar "Reusable Software Component Life Cycle" International Journal of Computers, Vol.5, Issue 2 (2011), pp. 191-199.
- [2] G. Singaravel, V. Palanisamy and A. Krishnan "Overview Analysis of Reusability Metrics in Software Development for Risk Reduction" International Conference on Innovative Computing Technologies (ICICTI), February, 12-13 (2010), Tamil Nadu, India.
- [3] R. Kamalraj, B. G. Geetha, and G.Singaravel "Reducing Efforts on Software Project Management using Software Package Reusability" IEEE International, Advance Computing Conference, March, 6-7 (2009), pp.1624-1627, Patiala, India,
- [4] P. S. Sandhu, Aashima, and P.Kakkar "A Survey on Software Reusability" 2nd International Conference on Mechanical and Electrical Technology (ICMET), September, 10-12 (2010)
- [5] S. Singh, M. Thapa, and S. Singh, et al. "Software Engineering, Survey of Reusability Based on Software Component" International Journal of Computer Applications (0975 – 8887) Vol. 8– No.12, October (2010), pp.39-42.
- [6] F. Amin, A. K. Mahmood, and A. Oxley "A Proposed Reusability Attribute Model for Aspect Oriented Software Product Line Components" Information Technology (ITSim), International Symposium, June, (2010).
- [7] G. Sindre E, A. Karlsson, and T. Staalhane "A Method for Software Reuse through Large Component Libraries" Fifth International Conference on Computing and Information. Proceedings ICCI 'May (1993).
- [8] W. J. Salamon and D. R. Wallace "Quality Characteristics and Metrics for Reusable Software" Preliminary Report, U.S. Department of Commerce, (1994).
- [9] K. K. Aggarwal, Y. Singh and A. Kaur, et al. "Software Reuse Metrics for Object-Oriented Systems" Proceedings of the 2005 Third ACIS Int'l Conference on Software Engineering Research, Management and Applications (SERA'05), (2005).
- [10] A. O. Bajeh. "A Novel Metric for Measuring The Readability of Software Source Code", Journal of Emerging Trends in Computing and Information Sciences. Vol. 2. No. 10. October (2011), pp.492-497.

Muhammad Ilyas received a Master degree in Software Project Management in 2004 from National University of Computer and Emerging Sciences, Lahore and a Doctor of Informatics from Johannes Kepler University, Linz Austria in 2010. His research interests include Software Engineering, Design Pattern and knowledge base systems. He is currently an assistant professor in the Department of Computer Science and Information Technology at the University of Sargodha, Pakistan.

Mubashir Abbas is a student of the Master of Science in Computer Sciences at the University of Sargodha. He has already completed his BS 4 year degree in computer sciences. His research interests include Software Engineering, Software Reusability and other topics.

Khansa Saleem is a student of the Master of Science in Computer Sciences at the University of Sargodha. Her research interests include Software Engineering, Semantic searches and other topics.

Sentiment Analysis of Equities using Data Mining Techniques and Visualizing the Trends

Shradha Tulankar¹, Dr Rahul Athale², Sandeep Bhujbal³

¹Department of Advanced Software and Computing Technologies
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India

²Sunflower Information Technologies Pvt. Ltd Pune, Maharashtra 411 016, India

³Department of Advanced Software and Computing Technologies
IGNOU – I²IT Centre of Excellence for Advanced Education and Research
Pune, Maharashtra 411 057, India

Abstract

Markets reflect sensitively towards opinions and sentiments. Investors should be wary of the fact that external factors interact deeply with the share markets and mark their influence over the coming period. National tragedy, low economic growth, IPO releases, RBI decisions on interest rates, foreign matters and many more have an increasing impact on public mindset. Even inflationary markets show a good impact on public opinions and also country's future growth prospects. There is no such technique to beat markets outcome, yet technical analysis and fundamentalist approach give you sustained results. In this paper, we have prepared a model that can predict the current market trends along with accuracy measures based on sentiment analysis. Sentiment analysis is an emerging trend to judge the market highs and lows. Visualization is provided to help end users understand markets properly and decide on their investment strategy. This model uses manual as well as automated sentiment analysis. Research reports are also taken as a mode of comparative assessment for the automated analysis.

Keywords: *Sentiment analysis, sentiments, visualization, predicts, markets, accuracy.*

1.0 Introduction

Investors invest a lot of time in checking out which strategies will be beneficial for them. Lot of research is conducted to explain the investor's behavior, his risk perceptions, market approach, trends of investing, factors affecting his investment and behavior towards the current market trends. Decisions made by a person solely depend on the statistics of the current market and whether the market will be bullish or bearish in coming days. This helps in deciding one's investment strategy [9].

There is a chaos of public sentiments and opinions on market fluctuations. This can be widely seen on chat forums, blogs, message boards or public forums. Facebook and Twitter are major source of such data. Ray Chen and Marius Lazer in their paper have devised a model to investigate the relationship between Twitter feed content and stock market movement [1].

Robert P. Schumaker, Yulei Zhang and Chun-Neng Huang have built and tested the financial news article system that incorporated sentiment analysis techniques in its predictive arsenal [2]. E. Bennet, M. Selvam, Eva Esther Shalin Ebenezer analyses individual investors sentiment. This study also analyses the influence of stock specific factors on investors' sentiment [3].

This project is based on designing an approach to perform sentiment analysis on stocks. The concept involves web-mining, text mining inclusion of sentiment and opinion mining and data mining techniques.

It is a combined approach to model a technique that will analyze the user's comments and live data accessed from web; regarding the fall and rise of the markets on a daily basis. The visualization provides the positive, negative and neutral trend of markets. It is a comparative approach based on the predictions given by brokers and sentiment analysis performed on the live streaming data.

The results obtained from sentiment analysis are compared by visualization techniques with the closing price of EOD of stocks. Comparative study helps users to identify whether to Buy, Sell or Hold the stock.

2.0 Sentiment Analysis on Indian Markets

It would certainly not be a wrong thing to say that Indian markets are emotionally driven. Events like natural calamity, national tragedy, sudden economic growth or even financial budget announcements hold the power to churn out the market scenario. E. Bennet, M. Selvam, N. Vivek, and Eva Esther Shalin who undertook a study on Indian investment business by accessing the retail investors have pointed out that it is not the markets that don't behave neatly but also the individual decision makers who don't behave in accordance with the tenets of expected utility theory [4].

There are many obvious reasons to raise the speculations of how sentiment analysis will help in judging the markets. One of being that there are more number of people who stick to daily trading as compared to long term investment. As research shows; when a particular stock price falls because of news regarding that particular company many investors are averse at selling it at loss. However, there are many who still wait in anticipation that the price will increase and their profits will be gained [9].

Sometimes, technical and fundamental analysis advises to buy particular stocks but for all no reasons the stock is sold in markets at a bulk basis. A thorough research if conducted will also help us in predicting insider trading. In a manner adopting professional sentiment analysis techniques will only help to solve such weird activities.

Various studies have been conducted till date in other countries to predict the current trends. However, the research is still going and the rate of accuracy is also increasing.

3.0 Objective

This thesis concentrates on constructing a model which proves that stocks vulnerability can be predicted using sentiment analysis. We have conducted this research for Infosys stock. Initially we have collected the research reports published by brokers; namely Angel Broking and ICICI Direct Securities for the stock Infosys and month November, 2012 [7]. These reports were manually analyzed. It is not necessary that every time the predictions given by firms have always helped people to earn capital met their goals for all months.

This thesis conducts a survey of messages collected from the message boards[7][8] and applying sentiment analysis on the posts. Sentiment analysis can be done on wide source of data. Data can be in the form of messages, blogs, news corpus, research reports, etc. In this paper, we have used web mining to extract the messages from the websites. The chats that are used for analysis get updated continuously and hence it is very difficult to keep a track of page numbers while performing web mining.

Lastly, we extract the terms that determine positive, negative and hold tendency from the WordList. We have extracted these words manually. Classification, one of the data mining techniques is applied on the words generated to determine the rate of accuracy.

4.0 Constructing the Model

In this section we provide a thorough idea of how the model is built and the algorithm used to test the accuracy of sentiment analysis.

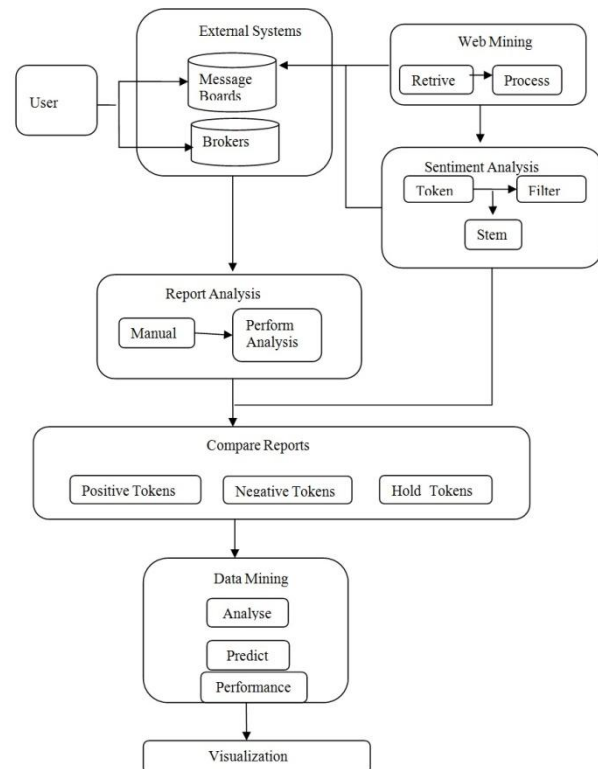


Fig 1: Proposed architecture of the model.

4.1 Data acquisition

The whole process is divided into three stages. Initially, the research is conducted on the data collected from the research reports. The research reports are thoroughly studied and verified manually.

Manual sentiment analysis is performed here to predict if the expectations have met their goals in real or not and if not then what are the specific reasons that went wrong in the analysis. Sentiment analysis is not bounded with the tools. It is mostly done manually and also with the help of tools. It cannot be restricted with a set of tools.

Secondly, we collect the data using web mining. Web mining allows you to look for patterns in data through content mining, structure mining, and usage mining. Content mining is used to examine data collected by search engines and web spiders. Structure mining is used to examine data related to the structure of a particular Web site and usage mining is used to examine data related to a particular user's browser as well as data gathered by forms the user may have submitted during Web transactions [13].

It is important to verify that the data meets the requirements of business problem. When doing predictive modeling the data also needs to contain the desired outcome. In this case the data was collected on following points:-

- Reviews given by people on the message boards.
- The comments given after trading hours were given higher preference.
- The broker's suggestion for a specific stock was collected.

4.2 Manual analysis of Research reports

Manual sentiment analysis is performed on the research reports collected online. Both the brokers gave the advice to buy Infosys stock for that month and hold it as a longterm investment.

4.3 Extracting contents from URL

Web structure mining, one of three categories of web mining for data, is a tool used to identify the relationship between Web pages linked by information or direct link connection. This structure data is discoverable by the provision of web structure schema through database techniques for Web pages [5].

This connection allows a search engine to pull data relating to a search query directly to the linking Web page from the Web site the content rests upon. Web mining allows the data to be structured and used for text mining [5].

The data acquisition happens in the form of web links, the data being in unstructured format. The content needs to be converted into documents for further processing. Data preprocessing is a mandatory task as the downloaded data contains all type of scripts, advertisements, codes and unnecessary words.

	A	B
1	links	Date
2	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
3	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/4/2012
4	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
5	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
6	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
7	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
8	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
9	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
10	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
11	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
12	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
13	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
14	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
15	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
16	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
17	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
18	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
19	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
20	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
21	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
22	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
23	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
24	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012
25	http://mb.moneycontrol.com/india/messageboard/view_topic_mspg.php?topic_id=4001&page=572&page=4&msg_type=8&int=3607&sq=290_381#top_pos	11/2/2012

Fig 2: ExampleSet of web links extracted according to date.

In this paper, we have made use of RapidMiner 5.0 along with web mining, text mining and Information Extraction plugins. RapidMiner 5.0 is an open source data mining tool and is widely used for text mining purposes [6].

For performing web mining the data needs to be imported. In this case the data had to be collected from dynamic website so accessing the contents using URL was the best method.

4.4 Generating WordList

The frequency with which particular words are used in a text can tell us something meaningful both about that text and also about its author because their choice of words is seldom random [15]. After extraction of URL, processing of unstructured data is performed and WordList is generated.

Text mining is performed using the Process document operator in Rapid Miner [14]. The links converted into documents are the input to the process document operator. Process document operator applies Tokens, Filter Stopwords, Stemming, generate n-grams and Transform Cases. This operator uses one single TextObject as input for generating a term vector [11].

Transform Cases is used to after converting the words into Tokens. This operator transforms all characters in a document to either lower case or upper case, respectively. This operator is then connected with Filter Stopwords (English). This operator filters English stopwords from a document by removing every token which equals a stopword from the built-in stopword list [11]. Many comments contain regional languages and hence this forms to be impurity for the processing of documents.

Filter Stopwords(English) is connected to N-Grams operator. This operator creates all possible n-Grams of each token in a document. A character n-Gram is defined as a series of characters of length n. The n-Grams of a token generated by this operator consist of all series of characters of this token which have length n. If a token is shorter than the specified length n, the token itself is kept in the resulting document [11].

After Filtering Stemming operator is used. Stemming operator splits the text of a document into a sequence of tokens. There are several options how to split the points; either we may use all non-letter character, which are the default settings. This will result in tokens consisting of one single word, which is the most appropriate option before finally building the word vector [11].

For Stemming purpose we use WordNet dictionary [16]. WordNet is a built in extension in RapidMiner. But with every installation we need to update this plugin also with other files. It is stored in the directory in a repository location of RapidMiner. Porter stemming is conducted on the WordList obtained after WordNet.

This operator stems English words using the Porter stemming algorithm applying an iterative, rule-based replacement of word suffixes intending to reduce the length of the words until a minimum length is reached [11].

The resulting WordList obtained after processing the documents contains 2500 records and generated for 18 models.

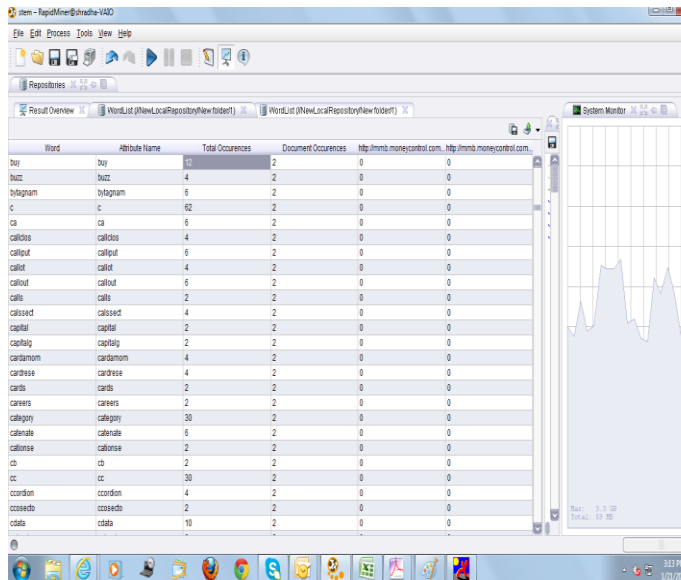


Fig 3:- WordList Generated for one trading day of 2500 records.

5.0 Predicting accuracy of model

Predictive analytics is the branch of data mining concerned with the prediction of future probabilities and trends. The central element of predictive analytics is the predictor, a variable that can be measured for an individual or other entity to predict future behavior. For predicting accuracy we used Weka, open source data mining tool [18].

Initially after automated analysis, from the WordList we extracted the words or phrases that showed the positive, negative or hold effects in market. This dataset was used for prediction.

The confidence measures are derived by applying Random Forest algorithm on the sheet. We obtained an accuracy of 87.5% for positive words. This indicates a strong possibility to buy the stock. The reports in turn were showing a high trend for buying the stocks and showing an upward trend.

6.0 Visualization

The visual data exploration techniques provide a much higher degree of confidence in the findings of the exploration. Visualization for this project was done in Mathematica 8.0 [12]. The visualization generally describes the impact of sentiments on market which is compared with the EOD closing price.

Following is the visualization report for positive, negative and hold terms. For generating WordList binary term occurrences was used instead of TF-IDF. Sometimes it is possible that a set of words will show the sentimental tendency and be more appropriate to judge the outcome of markets.

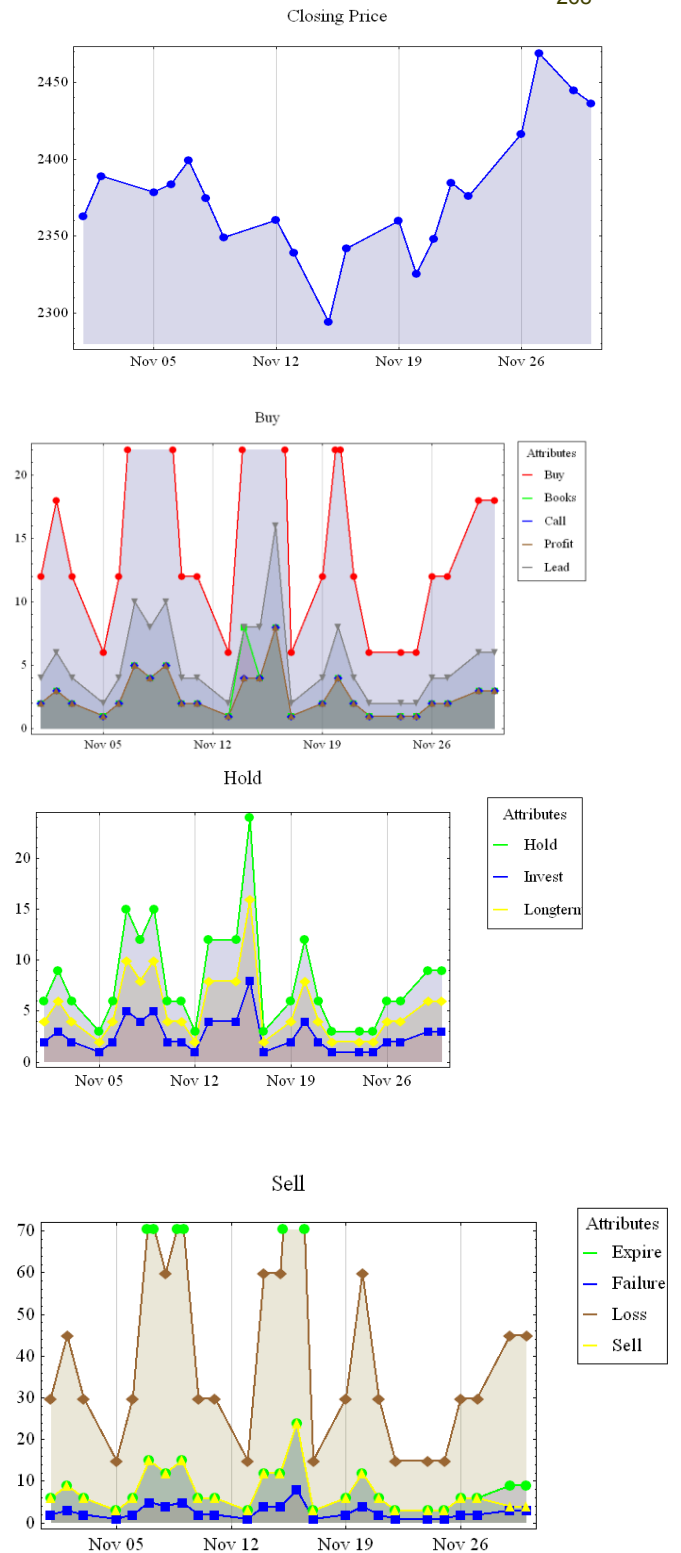


Fig 4:- Visualization graphs based on buy , sell and hold sentiment as compared with EOD closing price.

7.0 Conclusion

The stock has listed the highest number of document pages on 16th November 2012 and hence the profit during this day was highest. Similarly, on 23rd November 2012 it recorded very less documents. This accounted to the fact that the outside controversies prevailing during those days were

highest which had its impacts on Infosys stock during that tenure.

However, for comparative measures the predictions given by brokers are similar to those achieved by sentiment analysis. Hence it can be concluded that the stock had to be bought for that month.

Brokers had given accurate measures to buy the stock for the month of November 2012 and it will reach the level of 2400. The ending price of stock at 30th November 2012 clearly shows that it has recorded profit during that month and the stock can be recorded to positive trend.

Sentiment analysis is used for showing positive, negative and holds trends in the markets. The graphs when compared can be devised easily to show that the positive comments have higher trends in the graph.

In conclusion, the above model can be used for other stocks and hence it can be proved that sentiment analysis can be used to predict buy, hold or sell trends in the markets.

8.0 References

- [1] Ray Chen, Marius Lazer, "Sentiment Analysis of Twitter Feeds for the Prediction of Stock Market Movement".
- [2] Robert P. Schumaker, Yulei Zhang and Chun-Neng Huang, "Sentiment Analysis of Financial News Articles".
- [3] E. Bennet, M. Selvam, Eva Esther Shalin Ebenezer, "Investors' Sentiment influenced by Stock Specific Factors".
- [4] E. Bennet, M. Selvam, N. Vivek, and Eva Esther Shalin, "The impact of investors' sentiment on the equity market: Evidence from Indian stock market", 18 Oct. 2011.
- [5] Web DataMining, "Web Data Mining .net", [online] available at <http://www.web-datamining.net/content/>, 6 April 2013.
- [6] Wikipedia, "RAPID MINER", [online] available at <http://en.wikipedia.com/RapidMiner>, 15 Mar. 2013.
- [7] e-Eighteen.com Ltd, "moneycontrol.com", [online] available at http://mmb.moneycontrol.com/india/messageboard/view_topic_msgs.php?topic_id=4001&pgno=242&que=&msg_type=&tot=4603&pg_squen=141_461_#top_pos, 8 January 2013.
- [8] IRIS Business Services Limited, "myiris", [online] available at <http://myiris.com/newsCentre/storyShow.php?fileR=20121025111228199&dir=2012/10/25>, 12 January 2013.
- [9] Investopedia US, A Division of ValueClick, Inc., "Investopedia" "<http://www.investopedia.com/articles/05/032905.asp>", 05 March 2013.
- [10] Michael Wurst, Ingo Mierswa, "The Word Vector Tool and the RapidMiner Text Plugin", Version 5.2, March 14 2009.

- [11] Rapid Miner 5.2 User Manual, available online at http://docs.rapid-i.com/files/rapidminer/rapidminer-5.0-manual-english_v1.0.pdf, Copyright 2010.
- [12] Stephan Wolfram, "Mathematica Virtual Book", ISBN 1-57955-022-3, 2010.
- [13] Types of web mining, "<http://searchcrm.techtarget.com/definition/Web-mining>", 15 March 2013.
- [14] Vancouver Data Blogs for Text Mining; "<http://vancouverdata.blogspot.in/2010/11/text-analytics-with-rapidminer-loading.html>", 12 March 2013.
- [15] WordList ; "<http://www.ashgate.com/isbn/9780754672401>", 15 April 2013.
- [16] WordNet; "<http://wordnet.princeton.edu/>", 19 April 2013.
- [17] Predictive analytics; "<http://searchcrm.techtarget.com/definition/predictive-analytics>", 05 April 2013 .
- [18] Weka; "<http://www.cs.waikato.ac.nz/ml/weka/>", 20 April 2013.

First Author: She is pursuing **M.Tech** in Advanced Information Technology with specialization in Software Technologies from IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. She has pursued **B.E** Computer Science from University of Mumbai. Her research interests include Data Mining, Business Intelligence and Text Mining.

Second Author: He pursued **Ph.D.** from University of Linz Austria. He is Vice President **Sunflower Information Technologies Pvt. Ltd** Pune, India. He was faculty(Adjunct) in IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. His research interests include Mathematics, Business Intelligence, Parallel computing and Cloud computing.

Third Author: Sandeep Bhujbal is **Sr. Research Associate** in Advanced Software and Computing Technologies department of IGNOU – I²IT Centre of Excellence for Advanced Education and Research, Pune, India. He has pursued **M.C.S** from University of Pune. His research interests include Operating systems, Compiler construction, Programming languages and Cloud computing.

A Shadow Detection and Removal from a Single Image Using LAB Color Space

Ashraful Huq Suny¹ and Nasrin Hakim Mithila²

¹Department of Computer Science and Engineering, Chittagong University of Engineering and Technology
Chittagong 4349, Bangladesh

²Department of Computer Science and Engineering, Military Institute of Science and Technology
Dhaka-1216, Bangladesh

Abstract

Due to obstruction by an object light from a source cannot reach the area and creates shadow on that area. Shadows often introduce errors in the performance of computer vision algorithms, such as object detection and tracking. Thus shadow detection and removal is a pre-processing task in these fields. This paper proposes a simple method to detect and remove shadows from a single RGB image. A shadow detection method is selected on the basis of the mean value of RGB image in A and B planes of LAB equivalent of the image and shadow removal method is based on the identification of the amount of light impinging on a surface. The lightness of shadowed regions in an image is increased and then the color of that part of the surface is corrected so that it matches the lit part of the surface. The advantage of our method is that removing shadow does not affect the texture and all the details in the shadowed regions.

Keywords: Shadow detection, shadow removal, LAB color space, illumination.

1. Introduction

Shadows and shadings in images lead to undesirable problems on image analysis. That's why much attention was paid to the area of shadow detection and removal over the past decades and covered many specific applications such as traffic surveillance, face recognition and image segmentation. A shadow occurs when an object partially or totally occludes direct light from a source of illumination. In general, shadows can be divided into two major classes: self and cast shadows. A self shadow occurs in the portion of an object which is not illuminated by direct light. A cast shadow is the area projected by the object in the direction of direct light. Based on the intensity, the shadows are of two types – hard and soft shadows. The soft shadows retain the texture of the background surface, whereas the hard shadows are too dark and have little texture. Thus the detection of hard shadows is complicated as they may be mistaken as dark objects rather than shadows. Though most of the shadow detection methods need multiple images for camera calibration, the best technique must be able to extract shadows from a single image. This paper

gives a simple method to detect and remove shadows from a single RGB image.

2. Previous work

Removing shadows from images can significantly improve and facilitate the performance of certain computer vision tasks, such as tracking, segmentation, and object detection etc. It is therefore of great importance to discover ways of properly detecting shadows and removing them while keeping other details of the original image intact. Finlayson, Hordley and Drew [1] proposed a method which used illumination-invariant image with the original color image to locate the shadow edges. These edges are set to zero and the edge representation is reintegrated to get the shadow-free image. A faster method for shadow removal by averaging the results of reintegration along a few numbers of Hamiltonian paths in the image was proposed in [2]. Fredembach and Finlayson [3] proved that the error propagation during reintegration can be reduced by closing the shadow edges before reintegration.

In [4] the shadow removal is achieved in three stages. A First a 1D shadow-free illumination invariant image is created from which a 2D color representation is derived and then a 3D shadow-free color image is generated. The shadow edges are finally corrected by inpainting. Fredembach and Finlayson [5] suggested that the shadow regions differ from the non-shadow representation by a single constant which can be calculated in a little time. The constant for R, G and B channels are calculated separately. The constant is such that the addition of the shadow region with the constant will reduce the difference between the shadow region and the surroundings.

Xu, QI, and Jiang [6] proposed a method to detect vague shadows in an image using derivatives of the input image. A method to remove the shadows from curved areas retaining the background texture is proposed in [7]. The removal of shadows is achieved by calculating different scale factors for shadow regions and penumbra regions to cancel the effect of shadows.

Finlayson, Drew, and Lu [8] proposed that the shadows can be removed by minimizing entropy. An approach to

extract shadows from an image using the information supplied by the user is proposed in [9]. The image is segmented and the shadow, non-shadow and background regions are interactively specified by the user. The shadow removal is achieved by graph cut algorithm.

Zhu et al. [10] proposed a method to detect the shadows in single monochromatic image using a shadow invariant, shadow variant and near-black features.

A region-based approach to detect and remove the shadows from an image was proposed by Guo, Dai, and Hoiem [11]. The segmented regions in the image are classified based on relative illumination and using a graph-cut, the labeling of the shadow and non-shadow regions is done. The lighting of shadow-pixels is done to recover a shadow-free image.

A method to detect the shadows in a single image using a Tricolor Attenuation Model (TAM) was proposed in [12].

Salvador, Cavallaro, and Ebrahimi [13] proposed a method to identify and classify the shadows in color images. Luminance and color information are used to detect shadows. In [14] the shadow removal is done by illuminating the shadow region till it gets the same illumination as the surroundings. The texture is retained. The shadow removal is done using energy function in [15], assuming that the lighting needed in the shadow region is a constant.

Most of the works on shadow removal need multiple images and calibrated camera. Methods like reintegration using Poisson equation are time intensive. Also, dark objects are often mistaken as shadows. A simple method to detect and remove the shadows from a single RGB image is proposed in this paper.

3. Shadow Detection

To detect shadow initially the RGB image is converted to an LAB equivalent image. The LAB color space has three channels where L is the Lightness channel, A and B are the two color channels. The L channel has values ranging from 0 up to 100, which corresponds to different shades from black to white. The A channel has values ranging from -128 up to +127 and gives the red to green ratio. The B channel also has values ranging from -128 up to +127 and gives the yellow to blue ratio. Thus, a high value in A or B channel represents a color having more red or yellow and a low value represents a color having more green or blue.

Since the shadow regions are darker and less illuminated than the surroundings, it is easy to locate them in the L channel since the L channel gives lightness information. The B channel values are also lesser in the shadow areas in most of the outdoor images. Thus combining the values

from L and B channels, the pixels with values less than a threshold are identified as shadow pixels, and others as non-shadow pixels.

To detect shadow first a RGB image has to be converted to a LAB image. Then the mean values of the pixels in L, A and B planes of the image have to be computed separately. Now if $\text{mean}(A) + \text{mean}(B) \leq 256$, then the pixels with a value in $L \leq (\text{mean}(L) - \text{standard deviation}(L)/3)$ can be classified as shadow pixels and others as non-shadow pixels. Otherwise the pixels with lower values in both L and B planes can be classified as shadow pixels and others as non-shadow pixels [16].

This pixel-based method may classify some non shadow pixels as shadow pixels. Cleaning, a morphological operation can remove isolated pixels. The misclassified pixels are removed using dilation followed by erosion. Also area-based thresholding is done, so that only regions with a number of pixels greater than a threshold can be considered as shadow regions. All these morphological operations thus help to eliminate misclassification of pixels. Fig. 1 gives the shadow area detected using the proposed method.

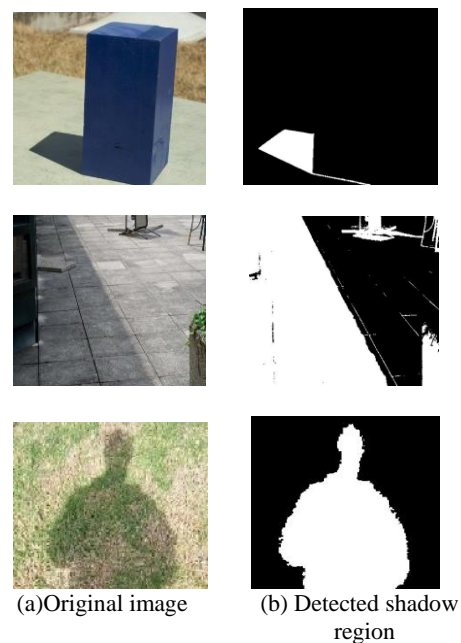


Fig. 1 Shadow area detected as white using the proposed method

4. Shadow Removal

The shadow removal technique used in [16] is done by multiplying the shadow regions by a constant and finally the shadow edge is corrected by using filters. The problem

with this approach is that dark objects may be considered as shadow and after removing the shadow some shadowed surfaces in the image still do not look similar to the non-shadow parts. This problem can be improved by the proposed method.

The intensity of light differs from the shadow region to the non-shadow region of an image. Actually the intensity of the shadowed regions gradually increases from shadow to light. We can say there may be two kinds of shadow region: partially lit region and non-lit region. So by lightening the shadow region can slightly remove the shadow from the image but not completely.

For complete removal of shadow we have to consider not only intensity but also hue and saturation of the shadow region. The chromaticity values of the image need to be corrected for this reason. In this paper, we use LAB color space and refer to a^* and b^* values in this color space as the chromaticity attributes. Applying a mean shift algorithm we segment the entire image according to its color values. Now the segments of the shadow region, P are certainly adjacent to a non-shadow region. Among all the neighbor segments, we choose the one that is closest in chromaticity to our segment of interest. This segment is further denoted as L. Afterwards, we rescale the shadowed segment's chromaticity values so that the average of the chromaticity in that segment P matches the average of the chromaticity in the aforesaid non-shadowed segment L.

$$a_p^* = a_p^* \frac{\langle a_L^* \rangle}{\langle a_p^* \rangle}$$

$$b_p^* = b_p^* \frac{\langle b_L^* \rangle}{\langle b_p^* \rangle}$$

Where a_p^* and b_p^* are a^* and b^* attributes of the corresponding segment and $\langle \bullet \rangle$ is the average operator. The chromaticity correction is valid for the surfaces that are partly lit and partly under shadow. For such regions, the chromaticity can be corrected so that the shadow part of the surface will have the same chromaticity as the non-shadow part of the surface.

Changing the chromaticity values of the surfaces which are completely under shadow to the closest adjacent non-shadow segment will introduce wrong colors. To prevent this effect, if the chromaticity difference between P and L is not small enough, the chromaticity value of segment P will not be changed. Using same process we continue to rescale the chromaticity values of non-lit regions. Finally, all boundaries between shadowed regions and neighboring non-shadow regions are smoothed by convolving them with a Gaussian mask. Thus, we introduce a uniform transition between shadowed regions that were lightened and neighboring non-shadowed regions.

4. Results and Discussion

Fig. 2 shows the result of applying our method on some images. The results are compared to the shadow removal method used in [16].

It can be clearly seen from column 'c' of Fig 2 that the texture of the surface that was under the shadow is preserved to a good extent and no harsh transition between the shadowed parts and non-shadowed parts can be seen. It is the main advantage of this paper.



Fig. 2 Column (a) shows the original image; Column (b) shows the shadow removed area using [16]; Column (c) shows the removed shadow region using proposed method

4. Conclusions

This paper describes a shadow removal method for real images based on increasing the lightness of shadowed regions in an image. The color of that part of the surface is then corrected so that it matches the lit part of the surface. Our algorithm worked successfully in removing both partially lit and non-lit regions.

References

- [1] Finlayson, G.D., S.D.Hordley, M.S.Drew, "Removing Shadows from Images", In: Proceedings of 7th European Conference on Computer Vision – Part IV, ECCV'02, London, UK, Springer-Verlag, 2002, 823-836.
 - [2] Fredembach, C., G.D.Finlayson, "Fast Re-Integration of Shadow Free Images", In: Proceedings of 12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications, CIC'2004, 9 November 2004, Scottsdale, Arizona, USA. IS&T – The Society for Imaging Science and Technology, 2004, 117-122.
 - [3] Fredembach, C., G.D.Finlayson, "Hamiltonian Path-Based Shadow Removal", In: Proceedings of 16th British Machine Vision.
 - [4] Finlayson, G.D., S.D.Hordley, C.Lu, M.S.Drew, "On the Removal of Shadows from Images", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 28, January 2006, No 1, 59-68.
 - [5] Fredembach, C., G.D.Finlayson, "Simple Shadow Removal", In: Proceedings of 18th International Conference on Pattern Recognition, ICPR'06. IEEE Computer Society, 2006, 832-835.
 - [6] Xu, L., F.Qi, R.Jiang, "Shadow Removal from a Single Image", In: Proceedings of 6th International Conference on Intelligent Systems Design and Applications, Vol. 2, 2006, 1049-1054.
 - [7] Arbel, E., H.Hel-Or, "Texture-Preserving Shadow Removal in Color Images Containing Curved Surfaces", In: CVPR IEEE Computer Society, 2007.
 - [8] Finlayson, G.D., M.S.Drew, C.Lu, "Entropy Minimization for Shadow Removal", Int. J. Comput. Vision, Vol. 85, October 2009, No 1, 35-57.
 - [9] Miyazaki, D., Y.Matsushita, K.Ikeuchi, "Interactive Shadow Removal from a Single Image Using Hierarchical Graph Cut", In: Proceedings of 9th Asian Conference on Computer Vision, Berlin, Heidelberg, 2010, Springer-Verlag, Volume Part I, ACCV'09, 234-245.
 - [10] Zhu, J., K.G.G.Samuel, S.Z.Masood, M.F.Tappen, "Learning to Recognize Shadows in Monochromatic Natural Images", In: CVPR, IEEE, 2010, 223-230.
 - [11] Guo, R., Q.Dai, D.Hoiem, "Single-Image Shadow Detection and Removal Using Paired Regions", In: CVPR, IEEE, 2011, 2033-2040.
 - [12] Rani, K.E.E., G.Jemilda, "Shadow Detection in Single Still Images Using TAM Based Multi Step Algorithm", IJCST, Vol. 2, 2011, Issue 4.
 - [13] Salvador, E., A.Cavallaro, T.Ebrahimi, "Shadow Identification and Classification Using Invariant Color Models", 2001.
 - [14] Germain, A., N.Salamati, S.Susstrunk, "Removing Shadows from Images Using Color and Near-Infrared", In: Proceedings of 18th IEEE International Conference on Image Processing ICIP'11, 2011, 1713-1716.
 - [15] Kumar, S., A.Kumar, "Shadow Detection and Removal in Color Images Using Matlab", International Journal of Engineering Science and Technology, Vol. 2, 2010, No 9, 4482-4486.
 - [16] Saritha Murali, V.K.Govindan, "Shadow Detection and Removal from a Single Image Using LAB Color Space", Cybernetics and Information Technologies, Vol. 13, 2013, No 1, Online ISSN: 1314-4081
- Ashrafal Huq Suny** will receive B.Sc. degree in Computer Science and Engineering (CSE) from Chittagong University of Engineering & Technology (CUET), Chittagong, Bangladesh in July, 2013. Currently he is a final year student Dept. of CSE. His research interest includes Digital Image Processing, Computer Vision.
- Nasrin Hakim Mithila** has received B.Sc. degree in Computer Science and Engineering (CSE) from Military Institute of Science and Technology (MIST), Dhaka, Bangladesh in 2012, December. Currently, she is employed as a Lecturer in Daffodil International University, Bangladesh and pursuing her M.Sc. from Bangladesh University of Engineering & Technology (BUET). Her research interest includes Wireless Networking and Digital Image Processing.

Software Reliability Prediction Using Artificial Techniques

Rita G. Al gargoor¹, Nada N. Saleem²

¹ Software Engineering Department, Mosul University, College of Computer sc. And Mathematics
Mosul, ZIP/+964, Iraq

² Software Engineering Department, Mosul University, College of Computer sc. And Mathematics
Mosul, ZIP/+964, Iraq

Abstract

Due to the growth in demand for software with high reliability and safety, software reliability prediction becomes more and more essential. Software reliability is a key part of software quality. Over the years, many software reliability models have been successfully utilized in practical software reliability engineering, however, no single model can obtain accurate prediction for all cases. So in order to improve the accuracy of software reliability prediction the proposed model combine the software reliability models with the neural networks (NN). Particle swarm optimization (PSO) algorithm has been chosen and applied for learning process to select the best architecture of the neural network. The applicability of the proposed model is demonstrated through three software failure data sets. The results show that the proposed model has good prediction capability and more applicable for software reliability prediction.

Keywords: *software reliability prediction, neural network, particle swarm optimization*

1. Introduction

As the use of software is increasing, the failures are also increasing rapidly. The consequences of failures may lead to loss of life or economic loss. So, the software professionals need to develop software systems which are not only functionally attractive but also safe and reliable [6]. Software reliability is defined as the probability of failure free software operation for a specified period of time in a specified environment[5]. Being able to predict the number of faults resides in software helps significantly in specifying/ computing the software release day and manage project resources (i.e. people and money) [23].

Software reliability growth models, refers to those models that try to predict software reliability from test data. These models try to show a relationship between fault detection data (i.e. test data) and known mathematical functions such as logarithmic or exponential functions. The goodness of fit of these models depends on the degree of correlation between the test data and the mathematical function[4]. Typically two broad categories of software reliability growth models (SRGMs) include parametric models and nonparametric models. Most of the parametric models are based on nonhomogeneous Poisson process (NHPP) that has been widely used successfully in practical

software reliability engineering[24]. Artificial neural network (ANN) with software reliability models have aroused more research interest .

in this paper, we use the effect of neural network to build non-parametric model for software reliability prediction, with the particle swarm optimization (PSO) algorithm used in our work for learning and to select the best architecture of the neural network.

The rest of this paper is organized as follows:

In section 2 a brief review of the researches carried out in the area of software reliability prediction is presented. Section 3 include background about software reliability. In section 4 the various artificial techniques that are applied in this paper are described briefly. Section 5 depicts neural network based approach for software reliability modeling. Section 6 presents the proposed model. Section 7 presents the experimental methodology and results. Section 8 conclude the paper.

2. Related Work

In recent years, many papers have presented various models for software reliability prediction . In this section, some works related to neural network techniques for software reliability modeling and prediction are presented.

Karunanithi et al. [18][16] first presented neural network based software reliability model to predict cumulative number of failures, the execution time is used as the input of the neural network. They used different networks like Feed Forward neural networks, recurrent neural networks like Jordan neural network and Elman neural network in their approach.

Karunanithi et al. [19] also used connectionist models for software reliability prediction, the results shows that the connectionist models may adapt well across different datasets and exhibit a better predictive accuracy.

Karunanithi et al. [17] they also predict the software reliability using neural network and present a solution to the scaling problem uses a clipped linear unit in the output layer of the neural network.

Su et al.[26] they used the neural network approach to build a dynamic weighted combinational model.

Sitte [25] presented a neural network based method for software reliability prediction. He compared the approach with recalibration for parametric models using some meaningful predictive measures with same datasets. They concluded that neural network approach is better predictors.

Cai et al. [8] proposed a neural network based method for software reliability prediction. They used back propagation algorithm for training. They used multiple recent 50 failure times as input to predict the next-failure time as output. They evaluated the performance of the approach by varying the number of input nodes and hidden nodes. They concluded that the effectiveness of the approach generally depends upon the nature of the handled data sets.

Hu et al. [12] proposed an artificial neural network model to improve the early reliability prediction for current projects/releases by reusing the failure data from past projects/releases.

Su et al. [27] proposed a dynamic weighted combinational model (DWCM) based on neural network for software reliability prediction. They used different activation functions in the hidden layer depending upon the software reliability growth models (SRGM).

Aljahdali et al. [2] investigated the performance of four different paradigms for software reliability prediction. They presented four paradigms like multi-layer perceptron neural network, radial-basis functions, Elman recurrent neural networks and a neuro-fuzzy model. They concluded that the adopted model has good predictive capability.

In [24], Singh et al. used feed forward neural network for software reliability prediction. They applied back propagation algorithm as learning algorithm. The experimental result had shown that the proposed system has better prediction than some traditional software reliability growth models.

Wang et al.[30] they used artificial neural network with software reliability combinational model by constructing the transfer function corresponding to the selected model.

In [13], Huang et al. derived software reliability growth models (SRGM) based on non-homogeneous poisson processes (NHPP) using a unified theory by incorporating the concept of multiple change-points into software reliability modeling. They estimated the parameters of their proposed models using three software failure data sets and compared results with some existing SRGM. Their model predicted the cumulative number of failures in various stages of software development and operation.

In [6][15][31], it was presented that the performance of a neural network system can be significantly improved by combining a number of neural networks.

3. Background

3.1 Concept of software reliability

Software reliability is one of the important factor been considered while ensuring the software quality. In simple term we can say that software reliability deals with the failure or faults that exist in the system [11]. Failures are the result of a fault in the software code, and several failures can be the result of one fault. The process of finding and removing faults to improve the software reliability can be described by a mathematical relationship called a software reliability growth model (SRGM)[3].

3.2 Software reliability growth models(SRGMs) and criteria

A software reliability growth model (abbreviated as SRGM) is known as one of the fundamental technologies for quantitative software reliability assessment, and playing an important role in software project management for producing a highly-reliable software system[14]. SRGM is mathematical model, shows how software reliability improves as faults are detected and repaired. SRGM can be used to predict when a particular level of reliability is likely to be attained. Thus, SRGM is used to determine when to stop testing to attain a given reliability level [21]. There are many software reliability growth models but the Commonly used model of software reliability models are JM, GO model, MO model, Sch model, S-Shape model. To evaluate the prediction powers of different models, it is necessary to use a meaningful measures. In this paper we use two criteria: Root Mean Square Error (RMSE) and Average Error(AE). These criteria are used to measure the difference between the actual and predicted values, the formulas is as Eq. (1) (2).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (c(k) - \hat{c}(k))^2} \quad (1)$$

$$AE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{c}(k) - c(k)}{c(k)} \right| \times 100 \quad (2)$$

Where n is the number of groups of failure data, $c(k)$ is the number of the actual failures in each group of failure data, $\hat{c}(k)$ is the number of the predicted failures. The smaller the RMSE and AE, the stronger that the model prediction ability[6] [30].

4. Overview of The Artificial Techniques Used

4.1 Artificial neural networks (ANNs)

An artificial neural network, or simply neural network, is a type of artificial intelligence (computer system) that

attempts to mimic the way the human brain processes and stores information. It works by creating connections between mathematical processing elements, called neurons, Fig.1 shows a neuron. Knowledge is encoded into the network through the strength of the connections between different neurons, called weights, and by creating groups, or layers, of neurons that work in parallel. The system learns through a process of determining the number of neurons or nodes and adjusting the weights for the connections based upon training data[27][29].

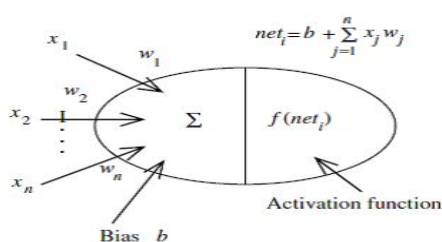


Fig. 1 The model of neuron

4.2 Particle swarm optimization(PSO)

Particle swarm optimization (PSO) is a population-based stochastic optimization technique modeled on the social behaviors observed in animals or insects, e.g., bird flocking, fish schooling, and animal herding. It was originally proposed by James Kennedy and Russell Eberhart in 1995. Since its inception, PSO has gained increasing popularity among researchers and practitioners as a robust and efficient technique for solving difficult optimization problems[7].

The swarm of particles initialized with a population of random candidate solutions move through the d-dimension problem space to search the new solutions. The fitness f_i can be calculated as the certain qualities measure. Each particle has a position represented by a position-vector present i (i is the index of the particle), and a velocity represented by a velocity-vector velocity i . After every iteration the best position-vector among the swarm is stored in a vector. The update of the velocity from the previous velocity to the new velocity is determined by Eq. (3). The new position is then determined by the sum of the previous position and the new velocity by Eq. (4).

$$\text{Velocity } ij(\text{new}) = w * \text{velocity } ij(\text{old}) + c1\text{rand1}(pbest \ ij(\text{old})) - \text{present } ij(\text{old}) + c2\text{rand2}(gbest \ j(\text{old}) - \text{present } ij(\text{old})) \quad (3)$$

$$\text{Present } ij(\text{new}) = \text{present } ij(\text{old}) + \text{velocity } ij(\text{new}) \quad (4)$$

Here w is the inertia weight, rand1 and rand2 are the random numbers usually chosen between $[0,1]$. $c1$ is a positive constant, called as coefficient of the self-recognition component, $c2$ is a positive constant, called as coefficient of the social component and the choice of value is $c1=c2=2$ generally referred to as learning factors[28].

4.3 Training the artificial neural network(ANN) using PSO algorithm

In this paper we use the particle swarm optimization(PSO) algorithm for learning and selecting the best architecture of our feedforward neural network, PSO is applied to feedforward neural network as follows:

The position of each particle in swarm represents a set of weights for the current epoch or iteration. The dimensionality of each particle is the number of weights associated with the network. The particle moves within the weight space attempting to minimize learning error. Changing the position means updating the weight of the network in order to reduce the error of the current epoch. In each epoch, the particles update their position by calculating the new velocity, and move to the new position. The new position is a set of new weights used to obtain the new error. This process is repeated and the particle with the lowest learning error is considered as the global best particle. The training process continues until satisfactory error is achieved by the best particle or computational limits (maximum iteration) are exceeded[1].

5. Neural Network Based Approaches for Software Reliability Modeling

5.1 The selection of the base models

As we mentioned in section 3 there are many software reliability models, among these models we choose GO model, logistic curve model, S-Shape model because their performance in software reliability evaluation. The failure mean value function of These three model are as Eq. (5)(6) and (7) [22].

$$m(t) = a(1 - e^{-bt}) \quad \text{GO model} \quad (5)$$

$$m(t) = \frac{a}{1 + ke^{-bt}} \quad \text{logistic curve model} \quad (6)$$

$$m(t) = a(1 - (1 + bt)e^{-bt}) \quad \text{S-Shape model} \quad (7)$$

5.2 Design the activation function of a neural Network for each base model

The derivation of the neural network into a software reliability modeling had depicted as follow:

For example, if we take the logistic growth curve model, this model simply fits the mean value with a form of the logistic function. so if we consider the basic feedforward neural network shown in Fig. 2, the derivation equations for the activation function of the neural network into software reliability models are depicted in Eq. (8),(9) and (10) .

$$y(t) = w_{11}^0 (1 - e^{-(w_{11}^1 t)}) \quad \text{GO model} \quad (8)$$

$$y(t) = \frac{w_{11}^0}{1 + e^{-(w_{11}^1 t)}} \quad \text{logistic curve model} \quad (9)$$

$$y(t) = w_{11}^0 (1 - (1 + w_{11}^1 t)) e^{-w_{11}^1 t} \quad \text{S-Shape model} \quad (10)$$

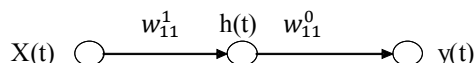


Fig. 2 Feed-forward neural network with single neuron in each layer

Where the input, $x(t)$, at time t is fed to the input layer, $h(t)$ is the output of the hidden layer, $y(t)$ is the output of the output layer[26].

6. Methods

In this section, we present the proposed model for software reliability prediction based on neural network.

6.1 Software reliability data

During the execution process the software may fail. Fig. 3 illustrate the software failure process, where t_i is the execution time for i th software failure and $\Delta t_i = t_i - t_{i-1}$ is the time interval between the $(i - 1)$ th and i th software failures [31]. The proposed model is used for predicting by taking execution time t_i as input and the accumulative number of failures N_i as output.

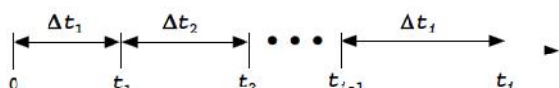


Fig. 3 Software failure process.

6.2 The proposed model

The prediction model based on neural network is shown in Fig.4, the input of the model is t_i which is encoded value of the execution time between (0,1), the output of the system is the cumulative number of failure N_i , the prediction model consist of three neural network combined together, each one is a three - layer feedforward neural network(FFNN) with n number of hidden neurons in the hidden layer, and there is a bias node in the input and the hidden layer, the activation function of the first neural network represent GO model, the activation function of the second neural network represent logistic curve model and the third neural network represent S-Shape model. For example the output of the first neural network as Eq. (11).

$$y(t) = 1 - e^{-((w_{21}(k1)+w_{22}(k2)+W_{23}(k3))+B*bw_{24})} \quad (11)$$

$$k1 = 1 - e^{-w_{11} t + B * b w_{11}}$$

$$k2 = 1 - e^{-w_{12} t + B * b w_{12}}$$

$$k3 = 1 - e^{-w_{13} t + B * b w_{13}}$$

the output of the three neural network are combined together using mean value rule which is defined as Eq. (12).

$$N'_i = \frac{1}{M} \sum_{m=1}^M N_{i,m} \quad (12)$$

Where $N_{i,m}$ is the output of 3 neural network, $m=1,2,\dots,M$.

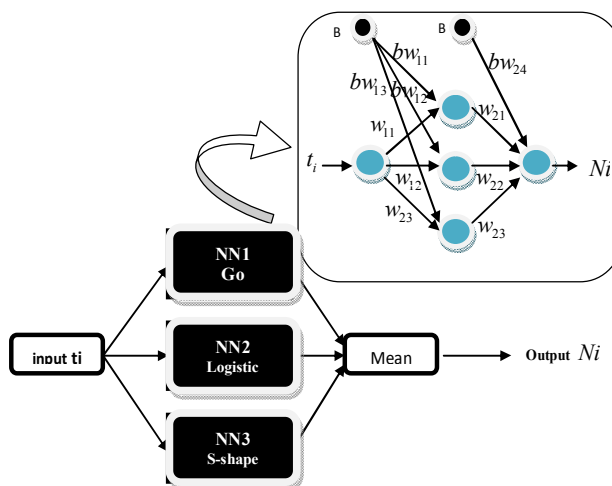


Fig.4 The architecture of the proposed model

6.3 The optimal set of PSO parameters

The PSO algorithm with well-selected parameter set can have good performance, In our software reliability prediction model, we use n number of swarms, the number of particle in each swarm is 25, the inertia weight is 0.9. according to Eberhart and Shi [10], the acceleration coefficients c1,c2 represent the stochastic acceleration that pulls each particle toward Gbest and Pbest position. In our model the acceleration coefficients c1,c2 changed in each iteration, and decreasing from large value to a small value through each iteration to improve the performance of the PSO algorithm, c1Max=0.9, c1Min=0.1, c2Max=0.2, c2Min=0.1.

6.4 Artificial neural network training and selecting the best architecture for the proposed model

The connection weights of the network have to be adjusted through a learning algorithm based on the training data, each neural network is trained with different initial weights connecting the three layers, the three NN trained in parallel by using particle swarm optimization(PSO) algorithm as we mentioned in section 4.

The number of neurons in the hidden layer determines the network's learning capabilities and its selection is a key issue in optimal network structure design. For selecting the best architecture of the NN(optimal number of the hidden layer neurons), we first build n models of NN architecture, each one have a different number of hidden neurons in the hidden layer, according to this we initialize n number of swarms based on the number of the models of the NN architecture. Fig.5 shows the flowchart of PSO learning process and selecting the best architecture. At the end of the training process (the maximum number of iterations), we get n number of Gbest (the optimal weights) of each swarm, then we evaluate the performance of each NN architecture based on the set of weights taken from Gbest, the evaluation is done by calculate the mean squared error(MSE), and the simulation results for the optimal number of the neurons in the hidden layer based on the minimum MSE.

7. Experiments

The proposed approach is applied to three software reliability datasets DS1[20], DS2[9] and DS3[23]. The DS1 and DS3 was collected from a real-time command and control application with 136 failures for DS1, 481 failures for DS3, the DS2 was collected from Operating System with 375 failures. The execution time and the number of failures of each dataset are normalized to the range of [0,1]. The model trained using 70% of the failure data for each dataset and the remaining data were used to

test the model, the evaluation criteria (AE, RMSE) of the training and testing are shown in table 1. Figures 6 to 11 are showing the training and testing result for various datasets using our proposed model.

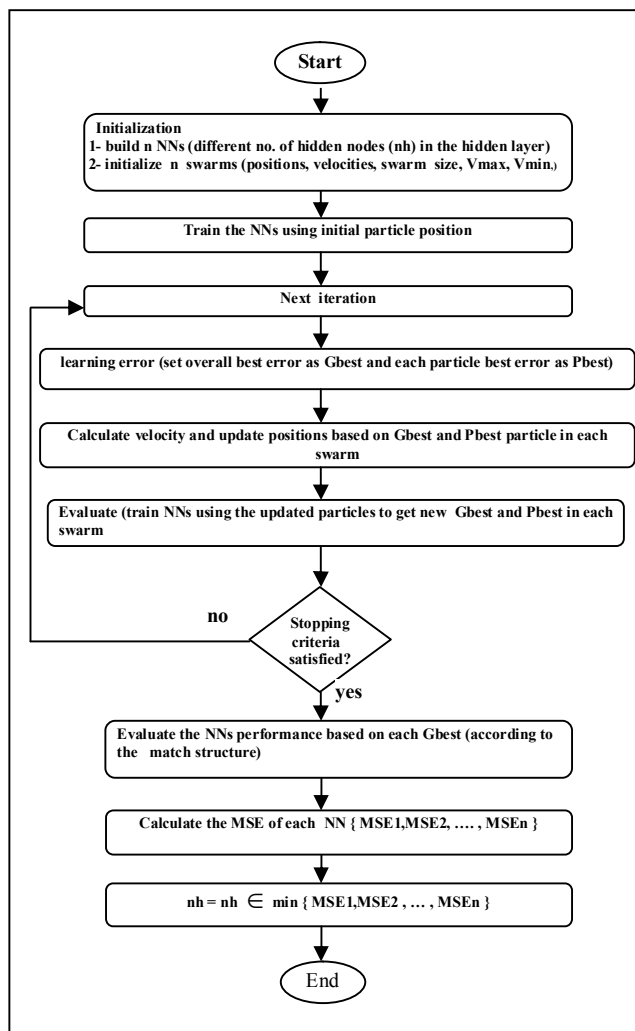


Fig. 5 the flowchart of PSO learning process and selecting the best architecture.

Table 1: results for the AE and RMSE of the training and testing dataset obtained using the proposed model

Dataset no.	Best no. of hidden neuron	Training data		Testing data	
		AE	RMSE	AE	RMSE
DS1	15	5.338824	0.016248	7.178521	0.018374
DS2	15	5.666911	0.016613	6.087315	0.018605
DS3	3	3.219852	0.013820	4.481290	0.018025

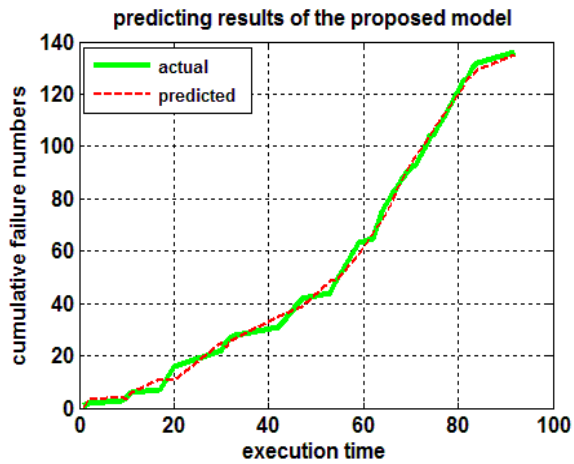


Fig. 6 Actual and predicted failure using DS1 (Training case)

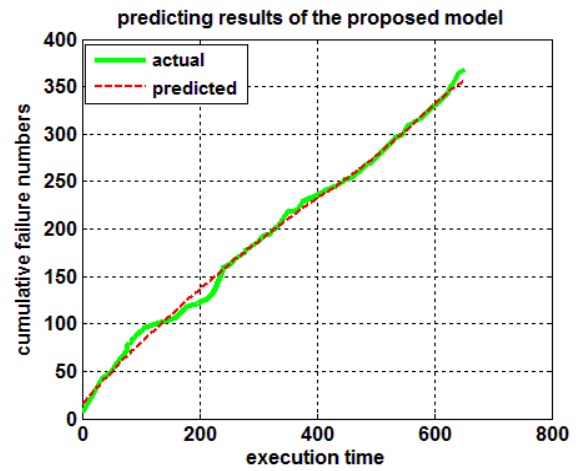


Fig. 9 Actual and predicted failure using DS2 (Testing case)

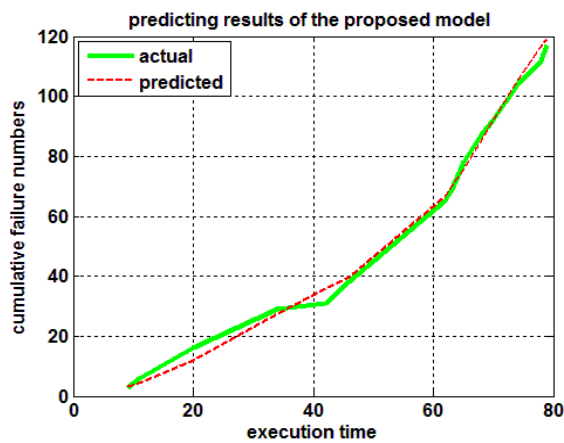


Fig. 7 Actual and predicted failure using DS1 (Testing case)

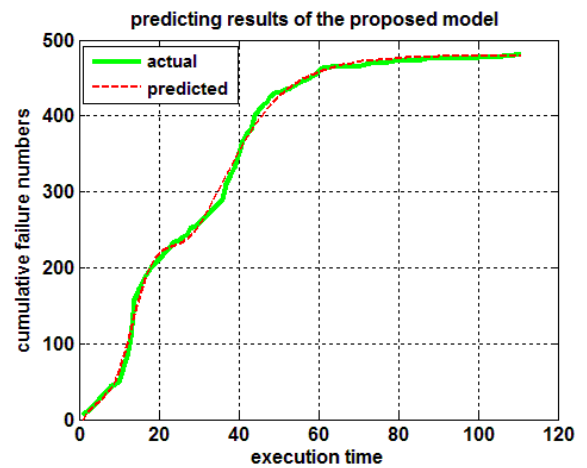


Fig. 10 Actual and predicted failure using DS3 (Training case)

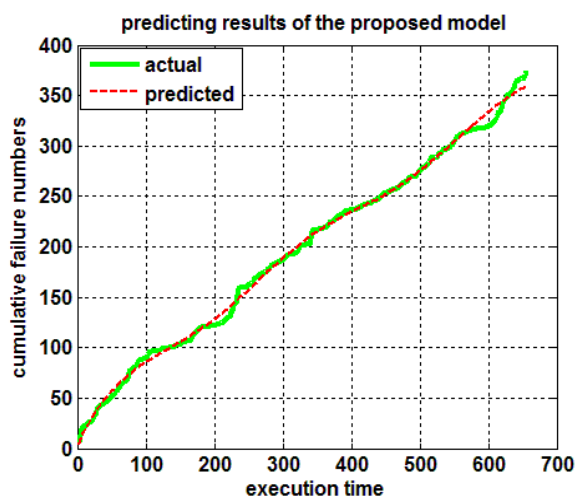


Fig. 8 Actual and predicted failure using DS2 (Training case)

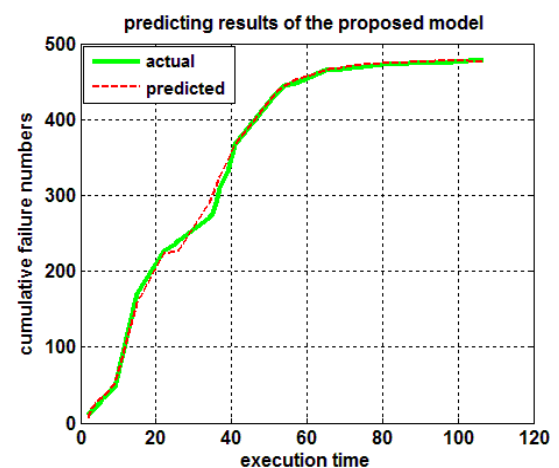


Fig. 11 Actual and predicted failure using DS3 (Testing case)

8. Conclusions

Software reliability is generally accepted as the major factor in software quality since it quantifies software failures. Neural networks trained by particle swarm optimization (PSO) has shown to be an effective non-parametric technique for software reliability prediction by optimizing the mean squared error(MSE), Selecting the best architecture of the network are also concerned for enhancing the performance of our model. The experimental results show that the proposed model gives acceptable result for different datasets relating to the prediction of the software cumulative failure.

References

- [1] Abdull hamed H. N., Shamsuddin S. M., and Salim N. , “ Particle Swarm Optimization for Neural Network Learning Enhancement” ,Jurnal Teknologi, 2008, pp.13 – 26.
- [2] Aljahdali S. H., Buragga K. A., “Employing four ANNs Paradigms for Software Reliability Prediction: an Analytical Study”, ICGST International Journal on Artificial Intelligence and Machine Learning, Vol. 8 , 2008, pp. 1- 8.
- [3] Almering V., Genuchten M. V. , and Cloudt G., “Using Software Reliability Growth Models in Practice” IEEE computer society, 2007, pp. 82 – 88.
- [4] Al-Rahamneh Z., Reyalat M. Sheta A. F., Bani-Ahmad S., and Al-Oqeili S., “A New Software Reliability Growth Model: Genetic-Programming-Based Approach” , Journal of Software Engineering and Applications, 2011, pp. 476-481.
- [5] Benaddy M., Wakrim M., and Aljahdali S., “evolutionary neural network prediction for cumulative failure modeling”, IEEE., 2009, pp. 179–184.
- [6] Bisi M. , and Goyal N. K., “ Software Reliability Prediction using Neural Network with Encoded Input ”, International Journal of Computer Applications , vol. 47, No. 22 , 2012, pp. 46 – 52.
- [7] Blum C. ,and Merkle D., Swarm Intelligence Introduction and Applications, Springer-Verlag Berlin Heidelberg., 2008.
- [8] Cai K.Y., Cai L. , Wang W. D., Yu Z. Y. ,and Zhang D., “ On the neural network approach in software reliability modeling” , The Journal of Systems and Software, vol. 58, 2001, pp. 47- 62.
- [9] A. Data & Analysis Centre for Software DACS , <https://www.thedacs.com/databases/sled>.
- [10] Eberhart R.C.,and Shi Y.,“ Particle Swarm Optimization : Developments, Applications and Resources” , IEEE, 2001 , pp. 81 – 86.
- [11] Haque F., and Bansal, S. “Software Reliability Estimation Models: A Comparative Analysis” , International Journal of Computer Applications, Vol. 43, 2012, pp. 27 – 31.
- [12] Hu Q.P., Xie M. and Ng S. H., “ Early Software Reliability Prediction with ANN Models” ,IEEE, 2006, pp. 210-220 .
- [13] Huang C. Y.,and Lyu M. R., “Estimation and Analysis of Some Generalized Multiple Change-Point Software Reliability Models” ,IEEE Transactions on Reliability, Vol. 60, 2011,pp. 498- 514.
- [14] Inoue S., and Yamada S.“Two-Dimensional Software Reliability Measurement Technologies”, IEEE , 2009, pp. 223 – 227.
- [15] Kapur P. K., Yadavalli V. S. S., Khatri S. K. , and Basirzadeh M. “ Enhancing Software Reliabilityof a Complex Software System Architecture Using Artificial Neural-Networks Ensemble” , International Journal of Reliability, Quality and Safety Engineering, Vol. 18, 2011, pp. 271–284.
- [16] Karunanithi N. , Whitley A., Malaiya Y. K. , “ Using Neural Networks in Reliability Prediction” ,IEEE , 1992, pp. 53 – 59.
- [17] Karunanithi N. ,and Malaiya Y. K., “The Scaling Problem in Neural Networks for Software Reliability Prediction” , IEEE , 1992, pp. 76 – 82.
- [18] Karunanithi N., Malaiya Y. K. ,and Whitley D. , “ prediction of software reliability using neural networks” , IEEE, 1991, pp. 124 – 130.
- [19] Karunanithi N., Whitley D., and Malaiya Y. K. , “ Prediction of Software Reliability Using Connectionist Models” , IEEE Transactions on Software Engineering, Vol. 18, 1992, pp. 563 – 574.
- [20] Lyu M. R., Handbook of software reliability engineering . New York: McGraw-Hill, 1996.
- [21] Quadri S. M., Ahmad N. ,and Farooq S. U. “ Software Reliability Growth Modeling With Generalized Exponential Testing –Effort And Optimal Software Release Policy” , Global Journal of Computer Science and Technology , 2011, pp.27 – 42.
- [22] Sharma K., Garg R., Nagpal C. K., and Garg R. K., “ Selection of Optimal Software Reliability Growth Models Using a Distance Based Approach” , IEEE TRANSACTIONS ON RELIABILITY, Vol. 59, 2010, pp. 266 – 276 .
- [23] Sheta A., “Reliability Growth Modeling for Software Fault Detection Using Particle Swarm Optimization” , IEEE Congress on Evolutionary Computation Sheraton Vancouver Wall Centre Hotel, 2006, pp. 3071 - 3078 .
- [24] Singh Y. ,and Kumar P., “Prediction of Software Reliability Using Feed Forward Neural Networks” , IEEE, 2010, pp. 1- 5.
- [25] Sitte R., “ Comparison of software–reliability-growth predictions:neural networks vs parametric recalibration,” IEEE Transactions on Reliability, Vol. 48, 1999, pp. 285- 291.
- [26] Su Y., Huang C., Chen Y., and Chen J. , “An Artificial Neural- Network-Based Approach to Software Reliability Assessment” , IEEE, 2005, pp. 1- 6.

- [27] Su Y., and Huang C. , “ Neural-network-based approaches for software reliability estimation using dynamic weighted combinational models” ,Elsevier Inc. The Journal of Systems and Software, 2007, pp. 606 – 615.
- [28] sumathi S. ,and surekha p. Computational Intelligence Paradigms Theory and Applications using MATLAB .Taylor and Francis Group, LLC., 2010.
- [29] Taylor B. J., Methods and Procedures For the Verification and Validation of Artificial Neural Networks, Springer Science + Business Media, Inc., 2006.
- [30] Wang G. ,and Li W., “ Research of Software Reliability Combination Model Based on neural Net” , IEEE, Second WRI World Congress on Software Engineering, 2010, pp. 253-256..
- [31] Zheng J. ,“Predicting software reliability with neural network ensembles” , Elsevier Ltd, 2009, pp. 2116 - 2122.

Spectral Subtraction for Speech Enhancement in Modulation Domain

Muhammad T. Sadiq¹, Noman Shabbir² and Wlodek J. Kulesza³

¹ Dept of Electrical Engineering, Sharif College of Engineering & Technology (Affiliated with UET Lahore),
Lahore, Pakistan

² Dept of Electrical Engineering, GC University,
Lahore, Pakistan

³ Dept of Electrical Engineering, Blekinge Institute of Technology,
Karlskrona, Sweden

Abstract

In this paper, an enhancement method is proposed for a speech signal that is corrupted by background noise. The proposed method is based on the spectral subtraction technique. The conventional spectral subtraction method improves the quality of the signal by reducing noise but it introduces an annoying musical noise. To eliminate the musical noise, spectral subtraction method in modulation domain is proposed. A filter bank, the heart of modulation domain, splits the broadband signal into several sub-bands. Then Instantaneous Frequency (IF) carrier estimator technique is used to estimate the carrier and modulating signal of each sub-band. Implementing spectral subtraction on each modulating signal and by synthesizing signals at the output, we get the enhanced speech signal without musical noise. Both qualitative (listening tests) & quantitative tests were performed to determinate the intelligibility and quality of speech enhanced by the proposed method.

Keywords: *Speech Enhancement, Spectral Subtraction, Modulation domain, Noise Removal.*

1. Introduction

SPEECH signals are low frequency modulating signals which modulate high frequency carriers. During transmission, these signals are affected by Additive White Gaussian Noise (AWGN). The quality of the speech signal is deteriorated by the noise recorded in real time situations. The noise decreases the intelligibility and quality of the signal while increasing listener fatigue. To enhance the speech and suppress the noise, different speech enhancement methods such as adaptive gain equalizer, Wiener filtering and spectral subtraction are used. Experimental results show that spectral subtraction reduces the background noise effectively but it introduces a musical noise in the signal [3]. The musical noise is an offensive noise. In order to eliminate the musical noise, the spectral

subtraction method is implemented in modulation domain. Modulation domain is used to split a speech signal into carrier and modulating signals and an Instantaneous Frequency (IF) technique estimates these signals.

Our proposed method not only enhances the speech signal but it also provides better background noise reduction than conventional spectral subtraction method and does not introduce musical noise in the signal.

2. Survey of Related Works

A spectral subtraction algorithm [1] suppresses stationary noise from speech by subtracting the spectral noise bias calculated during non-speech activities; however it attenuates the musical noise left after subtraction. The speech enhancement cascaded algorithm increases the overall effectiveness of the speech enhancement system [3]. Meanwhile the drawback of this algorithm is that it introduces undesirable musical noise. Modulation domain is a new domain in signal processing, in which one separates the modulating and carrier signals of the speech and different techniques like smoothed Hilbert carrier estimator, Instantaneous Frequency (IF) carrier estimator and frequency reassignment carrier estimator are used to estimate these signals [2]. Modulation domain processing has grown in popularity finding applications in areas such as speech coding, speech and speaker recognition, objective speech intelligibility evaluation as well as speech enhancement [5].

3. Problem Statement and Main Contribution

Speech signal is contaminated by background noise that appears randomly in the signal from frame to frame. This result in a twinkling sounding noise called musical noise that can be quite annoying for the listener.

Our research question is how to implement spectral subtraction speech enhancement method for musical noise elimination in a modulation domain. We hypothesized that the spectral subtraction method can be implemented in modulation domain using IF technique which should eliminate the musical noise.

The main contribution of this paper is to model and implement spectral subtraction in modulation domain using MATLAB and then to validate the method for speech enhancement.

4. Problem Solution

4.1 Modeling of Spectral Subtraction in modulation domain

Our proposed model is depicted in Fig. 1. In this model, a noisy signal $x(n)$ which is the combination of a clean speech signal $d(n)$ and a noise signal $v(n)$ is taken. A filter bank then divides the noisy speech signal into different sub-bands. The resultant signal is depicted as $x_N(n)$, where N is the number representing each sub-band. After that IF technique is used to estimate the carrier and modulating signal of each sub-band. The IF is the modification of the differential FM detector. A phase-only IF estimate $\alpha_N(n)$ is given by [2]:

$$\alpha_N(n) = \begin{cases} \left(\frac{Z_N(n)}{|Z_N(n)|} \right)^{\frac{1}{2}} & |Z_N(n)| > \epsilon \\ \alpha_N(n-1) & |Z_N(n)| \leq \epsilon \end{cases} \quad (1)$$

where $Z_N(n)$ is an un-normalized IF estimate, and ϵ is a small threshold that is used to reduce noise in the IF estimate. Smoothed IF estimates give an instantaneous phase estimate by the following equation, with the initial assumptions $W_N(-1)=1$ and $W_N(0)=\alpha_N(0)$:

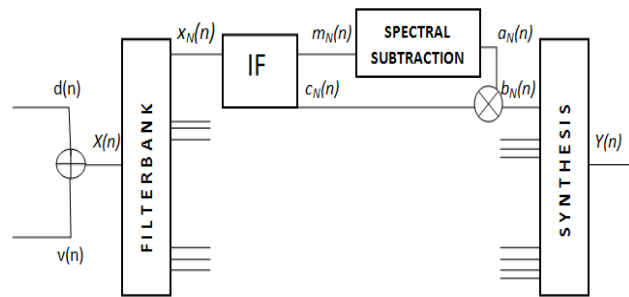


Fig 1: Block model of spectral subtraction in modulation domain.

$$W_N(n) = W_N(n-1)\alpha_N(n) \quad (2)$$

Once the instantaneous phase estimate $W_N(n)$ is calculated for sub-band $x_N(n)$, the carrier $c_N(n)$ is computed as:

$$c_N(n) = W_N(n) \quad (3)$$

and the modulating signal $m_N(n)$ can be found by demodulation of sub-band signal:

$$m_N(n) = x_N(n) \cdot c_N^*(n) \quad (4)$$

Now the spectral subtraction method is applied on the modulating signal $m_N(n)$. In spectral subtraction method, Voice Activity Detection (VAD) is used to determine the presence or absence of speech signal. Speech enhancer divides the input signal into different frames and then processes each frame. Let the signal of frame k -th be [4]:

$$m_{N,k}(n) = w(n)m_N(n+kD) \text{ for } n = 0 \dots L-1 \quad (5)$$

where k is the number of frame, D is the step size for each frame, $w(n)$ is a window and L is the length of a frame.

The power spectral density of noise for frame k is calculated as [4]:

$$P_{vv,N,k}(\omega) = \begin{cases} \gamma P_{mm,N,k}(\omega) + (1-\gamma)P_{vv,N,k-1}(\omega) & \text{if } k \text{ is nonspeech frame} \\ P_{vv,N,k-1}(\omega) & \text{if } k \text{ is speech frame} \end{cases} \quad (6)$$

where $P_{vv,N,k-1}(\omega)$ is the estimated power spectral density of noise for frame $k-1$, $P_{mm,N,k}(\omega)$ is the estimated power spectral density of modulating signal for frame k and $0 < \gamma < 1$ is a forgetting-factor.

To enhance the signal $m_N(n)$, a gain function $G_{N,k}(\omega)$ is define as:

$$G_{N,k}(\omega) = 1 - \frac{P_{vv,N,k}(\omega)}{P_{mm,N,k}(\omega)} \quad (7)$$

The output signal for k -th frame is calculated as:

$$A_{N,k}(\omega) = M_{N,k}(\omega) \cdot G_{N,k}(\omega) \quad (8)$$

where $M_{N,k}(\omega)$ is the spectrum of k^{th} frame of the modulating signal.

The output of spectral subtraction method $a_N(n)$ is obtained from $A_{N,k}(\omega)$ by using the overlap-add method. Each frame of $a_{N,k}(n)$ is calculated as the inverse Fast Fourier Transform (FFT) of $A_{N,k}(\omega)$. These frames are then rearranged after each other with respect to the frame length L and frame step size D . After that the signal $a_N(n)$ is calculated by adding the signal values where the frames overlap [4].

The signal $b_N(n)$ is the product of $a_N(n)$ and the carrier signal, $c_N(n)$. Similarly, this whole process is applied on each sub-band and after synthesizing signals at output we achieve enhanced speech signal $Y(n)$.

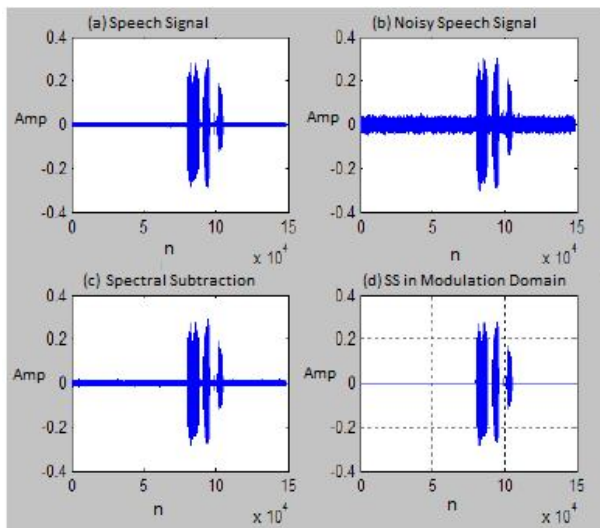


Fig 2: The enhancement effects.

4.2 Model Implementation and Results

Our model is implemented in MATLAB. For evaluation purposes a male's speech "The world is beautiful", recorded with little noise and the sampling frequency of 16 kHz is used. In Fig. 2(a), the original speech signal is depicted. The noise presented in the signal is not much so there is no need for enhancement at this stage. AWGN (with SNR=0dB) is added in the signal to make the scenario more practical. In Fig. 2(b), the speech signal contaminated with noise is represented. Then conventional spectral subtraction method is applied on this noisy speech signal and the output signal is shown in Fig. 2(c). Now our proposed spectral subtraction speech enhancement method in modulation domain is applied on the noisy speech signal and the result is shown in Fig. 2(d).

The validation of our model is firstly done by listening tests. Seven listeners participated in this experiment. The listening tests were conducted in a quiet room. Also in Fig. 2(c), the enhanced signal with conventional spectral subtraction is shown. It can be seen from the figure that it still suffers from undesirable musical noise. Musical noise is still present in the signal because modulating signal was not detected by using IF technique. But in our case, IF is used to detect modulating signal and then spectral subtraction method is applied in modulation domain. After listening tests, different listeners confirmed that our proposed method in modulation domain completely eliminates musical noise from the signal. This can also be verified by the signal represented in Fig. 2(d).

For quantitative purpose, the performance is measured through the Signal to Noise Ratio (SNR) improvement. The results were taken by varying number of sub bands. Numbers of sub-bands used are 2, 4, 8 and 16. In Fig. 3, SNR values for each band with conventional spectral subtraction method are shown. Improved values of SNR with our proposed method are depicted in Fig. 4. It can be seen from these figures that SNR value is around 12 dB for conventional spectral subtraction method where as spectral subtraction method in modulation domain provides SNR around 14 dB. There is an improvement of 2 dB in each sub-band's SNR.

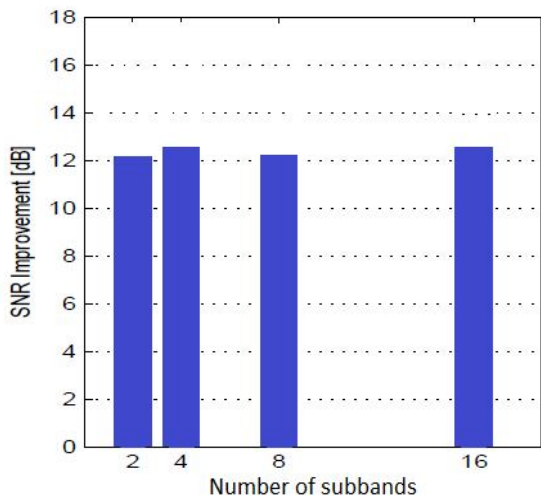


Fig 3: SNR without improvement

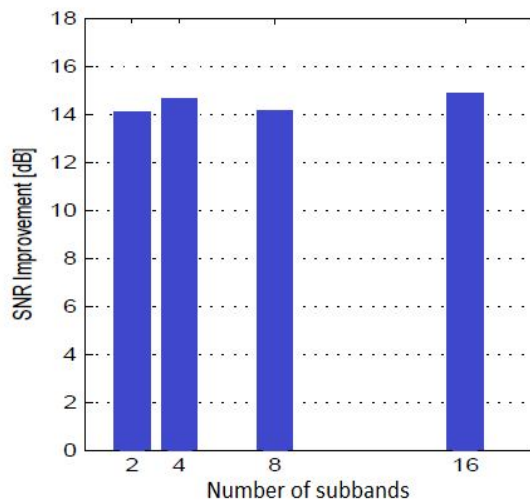


Fig 4: SNR after Improvement

4. Conclusions

In this paper, we proposed spectral subtraction method in modulation domain to enhance noisy speech with AWGN. Fig. 2(d) demonstrates that the modulation domain processing is useful when IF is used for the detection of modulating signal. Also the problem of musical noise is removed by using this technique.

The problem observed in the results is that a little amount of background noise is still present which requires to be improved. Though this noise does not effect much the intelligibility of speech but can be further improved.

References

- [1] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, pp. 113–120, 1979.
- [2] R. Ishaq, "Adaptive gain equalizer and modulation frequency domain for noise reduction," Master's thesis, Dept. Signal Processing, BTH, Karlskrona, Sweden, 2010.
- [3] T. K. Kumar, "Optimum cascaded design for speech enhancement using kalman filters", *WASET*, vol. 41, pp. 355-359, May 2008.
- [4] T. S. Gunawan, "Audio compression and speech enhancement using temporal masking models," Ph.D. dissertation, Dept. Elect. Eng., New South Wales Univ., Sydney, NSW, 2007.
- [5] L. Atlas, P. Clark, and S. M. Schimmel, "Modulation toolbox for matlab," Available: <http://isd1.ee.washington.edu/projects/modulationtoolbox/>



Muhammad Tariq Sadiq was born in Lahore, Pakistan, in 1985. He received his BS Electrical Engineering degree from COMSATS Institute of Information Technology (CIIT), Lahore, Pakistan, in 2009 and MS Electrical Engineering degree from Blekinge Institute of Technology, Sweden in 2011. He is currently working as a Lecturer & Project Manager in the Dept. of Electrical Engineering, Sharif College of

Engineering & Technology (Affiliated with UET Lahore) Lahore, Pakistan. His research interests are in the field of cognitive radio, speech processing and wireless networks.



Noman Shabbir was born in Lahore, Pakistan in 1985. He got his BS Computer Engineering degree from COMSATS Institute of Information Technology (CIIT), Lahore, Pakistan, in 2007 and MS Electrical Engineering (Radio Comm.) degree from Blekinge Institute of Technology, Sweden in 2009. His project on Unmanned Aerial Vehicles (UAV) got 3rd position in a National competition in 2006. He is currently working as a lecturer in the

Dept of Electrical Engineering, GC University, Lahore, Pakistan. His research interests are in the field of wireless networks and computer networks.



Wlodek J. Kulesza received the M.Sc. and the Ph.D. degrees from Lodz University of Technology, Poland, and a docent degree from Linköping University, Sweden. In 2001 he became Professor in Measurement Science at the University of Kalmar, Sweden. Since 2005 he has held a Professor position at the School of Engineering in the Blekinge Institute of Technology (BTH), Karlskrona, Sweden. His current research interests are multi-sensor systems and wireless sensor networks. IEEE member since 1995.

SPECTRUM HOLE PREDICTION BASED ON HISTORICAL DATA: A NEURAL NETWORK APPROACH

Bara'u Gafai Najashi^{1*}, Feng Wenjiang² Mohammed Dikko Almustapha³

^{1,2}College of Communication Engineering, Chongqing University,

Chongqing, China, Postal Code 400044

174 Shazheng Street, Shapingba District, Chongqing

³Department of Electrical and Computer Engineering, Ahmadu Bello University Zaria, Nigeria

ABSTRACT

The concept of cognitive radio pioneered by Mitola promises to change the future of wireless communication especially in the area of spectrum management. Currently, the command and control strategy employed in spectrum assignment is too rigid and needs to be reviewed. Recent studies have shown that assigned spectrum is underutilized spectrally and temporally. Cognitive radio provides a viable solution whereby licensed users can share the spectrum with unlicensed users opportunistically without causing interference. Unlicensed users must be able to sense whether the channel is busy or idle, failure to do so will lead to interference to the licensed user. In this paper, a neural network based prediction model for predicting the channel status using historical data obtained during a spectrum occupancy measurement is presented. Genetic algorithm is combined with LM BP for increasing the probability of obtaining the best weights thus optimizing the network. The results obtained indicate high prediction accuracy over all bands considered.

Keywords: Cognitive radio, spectrum sensing, prediction model, neural network, genetic algorithm

1. INTRODUCTION

The field of wireless communication has witnessed tremendous evolution over the years. The demand for bandwidth has grown with the introduction of several wireless standards. This advancement has brought about a perceived spectrum scarcity. Recent studies have shown that there is ample spectrum available that is not being utilized. So the problem can be said to be spectrum under utilization and not spectrum scarcity. Cognitive radio is seen as a great contender for better spectrum utilization because of its ability to make spectrum sharing possible. A situation whereby a licensed user is allowed to share the spectrum with an unlicensed user also called a secondary user opportunistically without causing interference will

surely increase the spectrum utilization. In order to achieve this, the unlicensed user must be able to sense when a channel is idle so that it can use, or busy so that it can vacate it. It has therefore become imperative to develop models capable of predicting the channel state accurately.

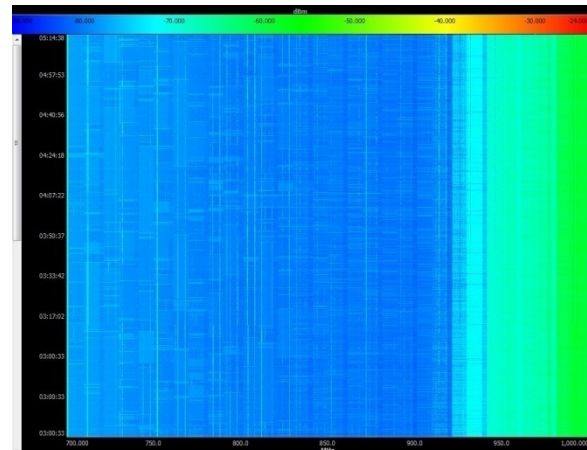


Figure 1: Waterfall figure showing sparsely used spectrum with low utilization level

Several spectrum prediction models have been proposed. Most of these works employ Markov chains for the prediction problem and assume that the primary user traffic follows a Poisson process [1][2][3]. In [4], a neural network based spectrum prediction using Multilayer Perceptron MLP was proposed. An hour long data was divided into 60 slots and converted into a time series. This was used to train and test the network. A practical spectrum behavior learning method based on MLP artificial neural network (ANN) was introduced in [5]. Performances were evaluated with an existing 7-days spectrum data set from a previous measurement, which was conducted in a metro city located in south

China. In this paper, a neural network based prediction model is presented. The training and test data were obtained from [10]. Unlike other neural network based prediction methods, a 12-hour long data was used for this work. The problem of weight selection which is common in neural networks was tackled using GA.

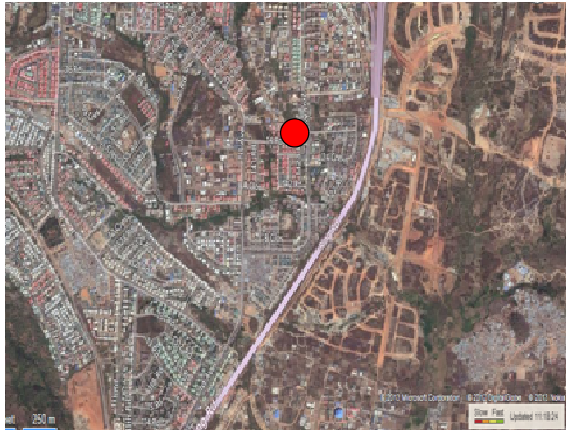


Figure 2: The red dot shows a bird's eye view of Gwarinpa, the measurement location in [10]

The rest of the paper is organized as follows. Section 2 provides a review of neural network, Multi-layer Perceptron and application of genetic algorithm in optimizing the interconnecting weights. In Section 3, the model used in this work is presented. Section 4 contains results obtained; conclusions are drawn from section 5.

2. NEURAL NETWORK

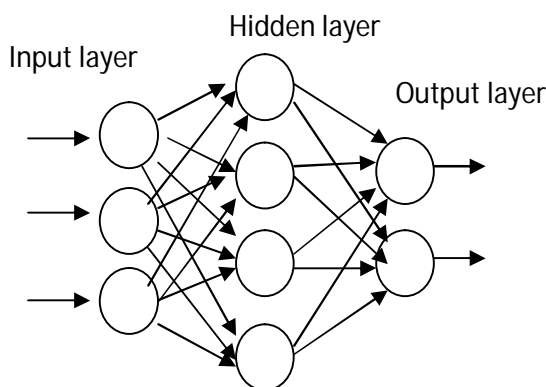


Figure 3: A basic neural network architecture showing the input, hidden and output layers

Artificial Neural Networks (ANNs) are non-linear mapping structures based on the function of the human brain. They are powerful tools for modeling and are widely used in prediction problems due to

their simplicity in terms of training. Whereas other prediction schemes require continuous training, neural network are trained once in an offline fashion when the observed process is stationary [6]. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. An ANN is configured for a specific application, through a learning process. Learning in biological system involves adjustment to the synaptic connections that exist between the neurons [7]. There are basically 2 types of architectures namely, the feed forward and the recurrent networks. The feed forward architecture is where the connections are feed forward this means that a neuron will not accept an input from a neuron it has previously feed. The recurrent networks are those that allow feedback between neurons. For this work the feed forward architecture is used.

2.1 Multilayered Perceptron (MLP)

An MLP network is a multi layered structure consisting of an input, output and a few hidden layers. Each layer (excluding the input layer) consists of a number of computing units called neurons which calculate a weighted sum of the inputs and perform a nonlinear transformation on the sum. The nonlinear transform is implemented using sigmoid functions (e.g. a hyperbolic tangent function). Neurons belonging to different layers are connected through adaptive weights. The output of a neuron j in the i^{th} layer, denoted by y_j^i , can be represented as:

$$y_j^i = \frac{1 - \exp(-v_j^i)}{1 + \exp(-v_j^i)} \quad (1)$$

Where $v_j^i = \sum_i y_i^{l-1} w_{ji}^l$ is the weighted sum of the inputs coming from the output of the neurons in the $(l-1)^{th}$ layer using the adaptive weights (or parameters) w_{ji}^l connecting the neuron j in the l^{th} layer and neuron i in the $(l-1)^{th}$ layer. Due to the nonlinear transform using the hyperbolic tangent function on v_j^i in (1), y_j^i lies in the range of $[-1, +1]$. If the inputs come from the input layer, v_j^i is calculated using the corresponding inputs instead of v_i^{l-1} . The total number of inputs in the input layer is referred to as the order of the MLP network and is denoted by τ . The number of hidden layers and the number of neurons in each layer depend on the application [8].

2.2 MLP Training

Once the architecture of a neural network has been determined for a particular application, the network is then ready for training. There are two types of training namely supervised and unsupervised training. In supervised learning, the input and output data is provided while in unsupervised learning the network has to make its decision on the input without any provision of the output. Over the years, MLP also known as a feed forward neural network has been applied successfully in several combinatorial problems. It works extremely well as a universal approximation in which input signal propagates in forward direction [11].

The general approach in learning involves presenting an input training pattern to an untrained network which is then used to get the output. The error is some scalar function of the weights that is minimized when the network outputs match the desired outputs. Since the actual output is known beforehand, the weights are adjusted to minimize the error so that the desired output could be reached. A Constant number 1/2 is added to do the mathematical derivation conveniently [5].

$$J(w) = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 \equiv \frac{1}{2} (t - z)^2 \quad (2)$$

Where t_k is the desired output and z_k is the actual output, t and z are the target network and output vectors with length c respectively. W represents the weights in the network. The initial weights are normally selected randomly and are updated in a way that reduces the error. This method is based on the gradient descent algorithm.

$$\Delta w = -\eta \frac{\partial J}{\partial w} \text{ or } (\Delta w_{pq} = -\eta \frac{\partial J}{\partial w_{pq}}) \quad (3)$$

η is the learning rate, which indicates the size of the change in weights. This iterative algorithm requires taking a weight vector at iteration m and updating it as

$$w(m+1) = w(m) + \Delta w(m) \quad (4)$$

Where m indexes the particular pattern presentation. Considering the hidden-output weights w_{kj} and applying chain rule differentiation

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}} = \partial_k \frac{\partial net_k}{\partial w_{kj}} \quad (5)$$

Where sensitivity of z_k can be defined to be

$$\partial_k \equiv -\frac{\partial J}{\partial net_k} = -\frac{\partial J}{\partial z_k} \frac{\partial z_k}{\partial net_k} = (t_k - z_k) f'(net_k) \quad (6)$$

Therefore, the weight update (learning rule) for the hidden-output weights given as

$$\Delta w_{kj} = \eta \partial_k y_j = \eta (t_k - z_k) f'(net_k) y_j \quad (7)$$

Similarly, considering the input-hidden weights w_{ji}

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \quad (8)$$

$$\frac{\partial J}{\partial y_j} = \frac{\partial}{\partial y_j} \left[\frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 \right] \quad (9)$$

$$= -\sum_{k=1}^c (t_k - z_k) \frac{\partial z_k}{\partial net_k} \frac{\partial net_k}{\partial y_j} \quad (10)$$

$$= -\sum_{k=1}^c (t_k - z_k) f'(net_k) w_{jk} \quad (11)$$

Thus the sensitivity of the hidden neuron can be given as

$$\partial_j \equiv \frac{\partial J}{\partial net_j} = -\sum_{k=1}^c \left(\frac{\partial J}{\partial z_k} \frac{\partial z_k}{\partial net_k} \frac{\partial net_k}{\partial y_j} \frac{\partial y_j}{\partial net_j} \right) \quad (12)$$

$$= -\sum_{k=1}^c \left(-\delta_k \frac{\partial net_k}{\partial y_j} \frac{\partial y_j}{\partial net_j} \right)$$

$$= f'(net_k) \sum_{k=1}^c (\delta_k w_{kj}) \quad (13)$$

Finally the learning rule for the input-hidden is given as

$$\Delta w_{kj} = \eta x_i \delta_j = \eta x_i f'(net_j) \sum_{k=1}^c (w_{kj} \delta_k) \quad (14)$$

By substituting equation (7) and (14) into equation (4), the updated weights could be obtained. The process will continue until the desired error θ goal is achieved.

$$J(W) = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 < \theta \quad (15)$$

2.3 Levenberg-Marquardt BP Algorithm

The Levenberg-Marquardt back propagation algorithm has been shown to be good at training moderately sized feed forward neural networks [12]. It is an approximation of Newton's method while back propagation with gradient descent technique is a steepest descent algorithm. It updates the weights using the expression below

$$\Delta w = - \left[\mu I + \sum_{p=1}^P J^p(w)^T J^p(w) \right]^{-1} \nabla E(w) \quad (16)$$

Where $J^p(w)$ is the Jacobian matrix of the error vector $e^p(w)$ evaluated in w , and I is the identity matrix. The error vector $e^p(w)$ is the error of the network for pattern p , that is

$$e^p(w) = t^p - o^p(w) \quad (17)$$

The parameter μ is increased or decreased at each step. If the error is reduced, then μ is divided by a factor β and multiplied by β in other case [13].

2.4 Genetic Algorithm

Genetic algorithms are stochastic search techniques that guide a population of solutions towards an optimum using the principles of evolution and natural genetics [7]. Genetic algorithms are inspired by the evolution of populations. In a particular environment, individuals who better fit the environment will be able to survive and hand down their chromosomes to their descendants, while less fit individuals will become extinct. The aim of genetic algorithms is to use simple representations to encode complex structures and simple operations to improve these structures. Genetic algorithms therefore are characterized by their representation and operators. In the original genetic algorithm an individual chromosome is represented by a binary string. The bits of each string are called genes and their varying values alleles. A group of individual chromosomes are called a population. [5]. Genetic algorithms are especially capable of handling problems in which the

objective function is discontinuous or non differentiable, non-convex, multimodal or noisy. Since the algorithms operate on a population instead of a single point in the search space, they climb many peaks in parallel and therefore reduce the probability of finding local minima [7]. Genetic algorithm involves 3 main operators: these include selection, cross-over (recombination) and mutation.

Selection is usually the first step in a genetic algorithm process. It involves selecting chromosomes from the population to crossover and produce offspring. The selection is based on Darwin's evolution theory "survival of the fittest" that the best chromosomes should survive and produce offspring. Extracting a subset of genes from a population according to a definition of quality (fitness function) is what the selection process is all about. Fitness function is the measure used to measure the optimality or closeness of the selected chromosomes to the desired ones. Many selection methods exist; they include Roulette wheel selection, Boltzmann selection, Tournament selection, rank selection, and steady state selection.

Crossover is usually the second stage. It's basically a process whereby two chromosomes combine to produce an offspring. The concept behind crossover is that an offspring may possess better qualities than its parents if the best qualities are transferred to the offspring. Crossover occurs during evolution according to a user-definable probability. Crossover selects genes from parent chromosomes and create a new offspring. Examples of crossover operators include one point crossover, two-point crossover, arithmetic and heuristic crossovers.

Mutation is an important step in genetic algorithm it helps in ensuring that the population is not stagnated at local minima. It is widely regarded as the last stage in the GA process. It helps in ensuring generic diversity from generation of population to the next. The process involves altering the gene values of one chromosome or even more. This can lead to the addition of totally new genes in the gene pools which will subsequently lead to a better result than initially obtained. Commonly used mutation operators include Flip-bit, boundary, non-uniform, uniform and Gaussian [14].

2.5 Optimizing BP-ANN with GA

Two factors generally influence modeling a network during the learning and training session. One is the initial interconnecting weights of the network, and another is their modified quantities [9]. Generally the

initial interconnecting weights of BP ANN are often stochastically and blindly produced, this might cause the network to run into partial optimization and therefore decrease the probability to obtain the optimal solutions. Moreover, because the Delta rule is always adopted to modify the interconnecting weights of BP ANN, the convergence velocity is always slow, or sometimes the network does not even converge. These shortages of BP ANN are quite necessary to be optimized and improved [9]. The problem of partial minimum of a BP ANN can be solved by adjusting the initial interconnecting weights of the network. This can be achieved through the application of Genetic algorithm because the problem is a non linear problem. GA is a nonlinear optimization method that has very strong ability of global searching.

weights and thresholds of BP ANN is considered to be an individual of the population and coded to be a chromosome. With the errors of BP ANN as the adaptive function, all the chromosomes are changed using GA operators (selection, crossover and mutation). The chromosomes that correspond to the best initial interconnecting weights and thresholds of BP ANN are gradually evolved. Finally, the initial interconnecting weights and thresholds optimized with GA are used into the learning and training of BP ANN again, and a better BP ANN model can be constructed. Because the initial interconnecting weights and thresholds of BP ANN are globally optimized by utilizing GA, an appropriate search space can be located in the complex solution space, and the partial minimum or non-convergence can be avoided.

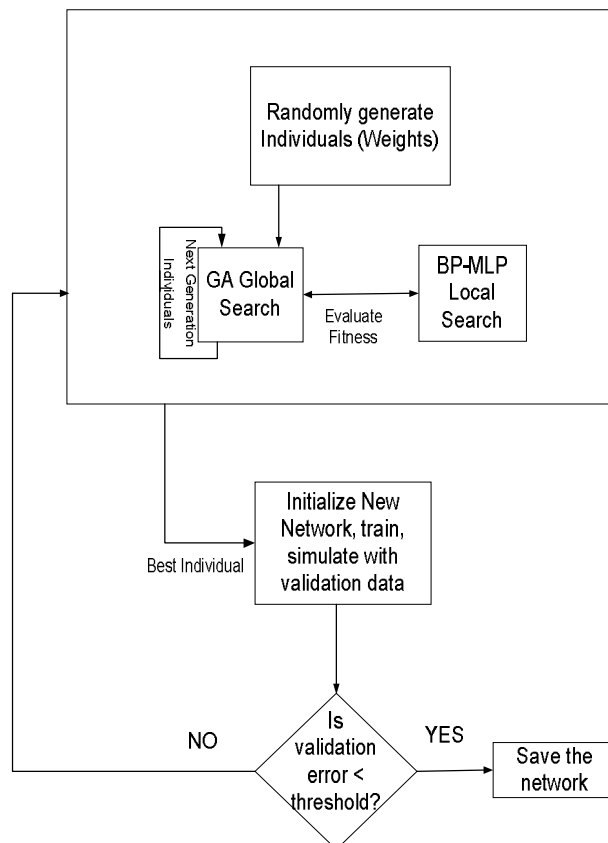


Figure 4: The flowchart of BP-MLP weight optimization using GA

After the structure and parameters of an ANN such as the number of layers and the number of neurons in every layer are determined, the approximate ranges of interconnecting weights and thresholds can be computed by using the algorithm of back propagation. Then, an initial population in the field of solution can be randomly produced by adopting genetic algorithm. Each group of the interconnecting

3. MODEL IMPLEMENTATION

Based on the 12 hour spectrum occupancy measurement conducted in [10], a total of 2,700 samples were obtained since each complete sweep takes 16 seconds. These raw power level data were processed and feed to the neural network. The input and output data were randomly selected from three different services namely: GSM band, broadcasting band, and the 3G cellular band. The input is a set of power level values obtained from the spectrum analyzer during the measurement for a single channel. In [10], the channel size was given as 200 KHz. The binary hypothesis defined as $X_t = 1$ if $X_t \geq m$, otherwise $X_t = 0$, where X_t are the power values from the spectrum analyzer and m is the decision threshold which was applied to get output data. The training and testing was achieved using data from the 875 MHz broadcasting band, GSM downlink 905 MHz, 3G downlink 1865 MHz and GSM uplink 890 MHz band. The entire process can be summarized in the following steps:

SERVICES	FREQUENCY
Broadcasting	700-806 MHz
GSM Uplink	890-895 MHz
GSM Downlink	905-910 MHz
3G 1800 Downlink	1865-1880 MHz
3G 1900 Downlink	1883-1890 MHz

Table 1: A table of different services considered

Step 1: The time series power level data is converted into a binary time series of 0's and 1's using thresholding.

Step 2: The data from step one is then grouped into the different services examined. The training and testing data sets are then obtained.

Step 3: The training and testing data for the service to be predicted is selected, architecture and training parameters are also selected.

Step 4: The network is trained using MLP-BP and weight selection is optimized through the use of GA.

Step 5: The network is tested with the test data to ascertain the prediction accuracy.

The simulation was conducted using MATLAB version r2012a. After a trial and error process, the number of neurons was set at 10 which were found to be adequate. GA and MLP- BP were used during the learning process and the prediction accuracy was determined across the bands considered.

Parameter	Value
Number of Hidden Neurons	10
Training samples	2700
Testing samples	2700
Transfer Function	Sigmoid
Training Algorithm	LM+GA

Table 2: Parameters used in modeling

4. RESULTS AND DISCUSSION

In this section, the performance of the designed neural network is given. The main attraction of using neural network based spectrum prediction lies in the fact that cognitive radios can learn and train itself from historical information obtained by the cognitive radio without redesigning the whole system completely as is the case with other methods. Unlike other models such as the Markov chain approach to spectrum prediction, the neural network model need only to be updated with the most recent data as its input. This approach saves power, sensing time and manpower. Machine learning has already been proposed as an integral part of future cognitive radios, the high prediction accuracy realized in this model will greatly reduce the time required to sense whole bands. In addition, the processing power required at the base station will be also reduced. The high prediction accuracy will also drastically reduce the rate of interference between the primary and secondary users in a cognitive radio scenario. It has therefore become paramount to improve machine learning based spectrum hole prediction accuracy in order to attain high prediction accuracy. For the purpose of this work, five popular services were considered. The GSM 900 uplink channels licensed to Etisalat had a mean prediction error of about 0.035

over five channels that were selected randomly. We observed a mean prediction error of about 0.005 in the GSM 900 downlink band. 0.0179 prediction error was recorded in the 3G downlink band. 0.0004 and 0.0007 were observed in the broadcasting band and 3G downlink band (licensed to Starcomms Nigeria) respectively. The difference between the uplink and downlink prediction accuracy can be due to the fact that the base station is continuously transmitting information to the mobile users which is not the case with the uplink band. It has already been stated in [5] that the sparsely used uplink band can provide an inspiration for future deployment of some form of cognitive radio in the future. The similarity in both GSM and 3G downlink bands can be due to the similar nature of the bands in terms of their behavior. The 1900 downlink band had an error of around 0.0007 which is very low. This value might be deceiving. This band is currently being used by Starcomms Nigeria which uses a CDMA based technology for their services. The spectrum analyzer might not be able to accurately detect the presence or absence of a signal during the spectrum occupancy measurement, therefore the low value experienced might be misleading because of the low power nature of the signals in this band. In a worst case scenario whereby a signal suffers from wireless effects such shadowing, multipath, and attenuation the power considerably degrades making detection almost impossible. Overall, the performance of neural network based spectrum prediction using GA for weight optimization is good.

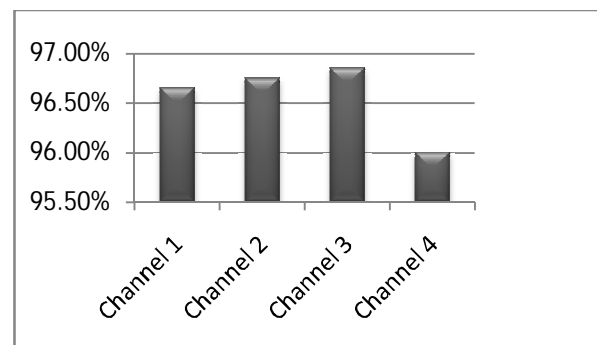


Figure 5: Spectrum Prediction Accuracy for 900 GSM Uplink

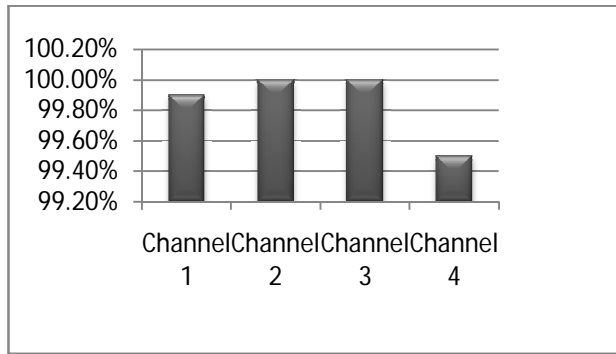


Figure 6: Spectrum Prediction Accuracy for 900 GSM Downlink

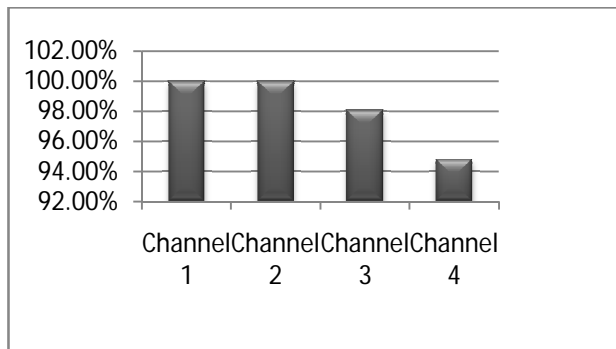


Figure 7: Spectrum Prediction Accuracy for 3G 1800 Downlink

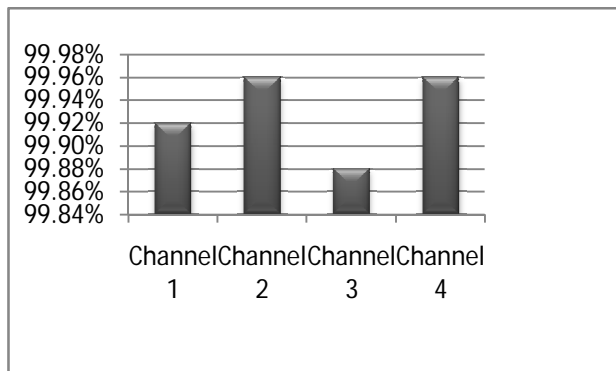


Figure 8: Spectrum Prediction Accuracy for broadcasting band

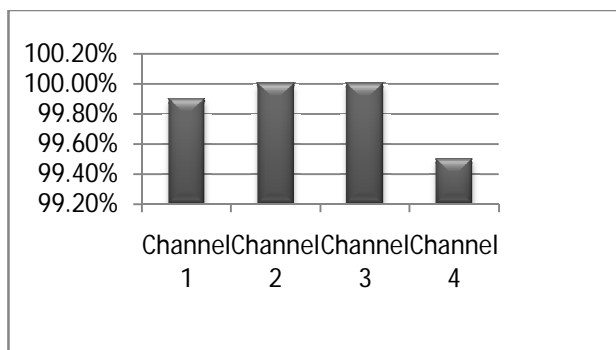


Figure 9: Spectrum Prediction Accuracy for 1900 3G Downlink

SPECTRUM BAND	PREDICTION ERROR
GSM 900 UPLINK	0.035
GSM 900 DOWNLINK	0.005
3G 1900 DOWNLINK	0.0179
BROADCASTING	0.0004
3G 1800 DOWNLINK	0.0007

Table 1: Prediction error for the services investigated

5. CONCLUSION

The importance of spectrum sensing is key to the development and eventual implementation of cognitive radios. Due to the random nature of spectrum which varies from one place to another, there is need to better understand this scarce resource so that its behavior can be predicted with little or no error. This knowledge could be obtained through extensive measurements that can help in developing models capable of predicting this behavior. Ultimately, these models could help in reducing sensing time and power consumption in cognitive radios. Unlike other models, neural network based models need not to be built from scratch once the model has been designed. Since it has already been suggested that cognitive radio will utilize a geo-location database which it will use to query information with respect to its current environment, machine learning based method could prove invaluable because of its ability to update itself whenever new information is available. Neural network based spectrum prediction has already been explored as was stated earlier; the aim of this paper was to explore the possibility of optimizing the initial weights in a neural network because they tend to run into partial optimization therefore reducing the accuracy of the deigned model. In this paper, a neural network based spectrum prediction model is presented. Unlike other models, the issue of local minimum in selecting interconnecting weights was addressed by using Genetic algorithm. Results indicate high prediction accuracy across all bands considered.

Acknowledgement

Our thanks go to the College of Communication Engineering, Chongqing University for their support and understanding during this work.

REFERENCES

1. I. A. Akbar and W. H. Tranter, "Dynamic spectrum allocation in cognitive radio using hidden Markov models: Poisson distributed case," in Proc. IEEE Southeast Con, 2007, pp. 196–201
2. C.-H. Park, S.-W. Kim, S.-M. Lim, and M.-S. Song, "HMM based channel status predictor for cognitive radio," in Proc. 2007 Asia-Pacific Microwave Conference (APMC), Dec. 2007, pp. 1–4.
3. V. K. Tumuluru, P. Wang, and D. Niyato, "Channel status prediction for cognitive radio networks," Wireless Commun. and Mobile Comput, Aug. 2010.
4. Zhao Jianli, Wang Mingwei, and Yuan Jinsha, "Based on neural network spectrum prediction of Cognitive Radio" Electronics, Communications and Control International Conference(IECC), 2011
5. Liang Yin, SiXing Yin, Weijun Hong, ShuFang Li, "Spectrum Behavior Learning in Cognitive Radio Based on Artificial Neural Network" Military Communication Conference, Track 1, Waveforms and Signal Processing, 2011.
6. Vamsi Krishna Tumuluru, Ping Wang and Dusit Niyato, "A Neural Network Based Spectrum Prediction Scheme for Cognitive Radio" IEEE International Conference on Communication (ICC), 2010.
7. Ms. Dharmistha, D. Vishwakarma, "Genetic Algorithm based Weights Optimization of Artificial Neural Network", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, August 2012.
8. Vamsi Krishna Tumuluru, Ping Wang and Dusit Niyato," A Neural Network Based Spectrum Prediction Scheme for Cognitive Radio",
9. Qiuwen Zhang, Cheng Wang," Using Genetic Algorithm to Optimize Artificial Neural Network: A Case Study on Earthquake Prediction", International Conference on Genetic and Evolutionary Computing
10. Bara'u Gafai Najashi,"Insight into Spectrum Occupancy in Nigeria", International Journal of Computer Science Issues", Vol 10, January 2013.
11. Shah, Habib, Rozaida Ghazali, and Nazri Mohd Nawi. "Using artificial bee colony algorithm for MLP training on earthquake time series data prediction. " *arXiv preprint arXiv:1112.4628* (2011).
12. Kuryati Kipli, Mohd Saufee Muhammad, Sh. Masniah Wan Masra, Nurdiani Zamhari, Kasumawati Lias, Dayang Azra Awang Mat," Performance of Levenberg-Marquardt Backpropagation for Full Reference Hybrid Image Quality Metrics", Proceedings of International Conference of Multi-Conference of Engineers and Computer Scientists", March 2012.
13. Alba, Enrique, and J. Chicano. "Training neural networks with GA hybrid algorithms." *Genetic and Evolutionary Computation-GECCO 2004*. Springer Berlin/Heidelberg, 2004.
14. Goldberg, D. E.: Genetic Algorithms in Search, Optimization, and Machine Learning [M]. Addison- Wesley Pub. Co., 1989.

The Effect of Data Hiding at Various Bit Positions on Audio Steganography in DCT Domain

Sahib Khan¹, Umar Said², Ejaz Ahmad³, Fazlullah⁴ and Mukhtair Ali⁵

¹⁻⁵ Dept. of Electrical Engineering, University of Engineering and Technology, Peshawar, Pakistan

Abstract

This work is related to audio Steganography and presents the statistical analysis of various quality measuring parameters of audio Steganography, e.g. signal to noise ratio (SNR), peak signal to noise ratio (PSNR) and mean square error (MSE), for different bit position substitution in an audio cover file in discrete cosine transform (DCT) domain. Interestingly the resulted mean square error (MSE), signal to noise ratio (SNR) and peak signal to noise ratio (PSNR) for each bit position shows a very insignificant variations in the values of the quality measuring.

Keywords: Audio Steganography, Discrete Cosine Transform, Signal Processing, Information Security

1. Introduction

Steganography is a technique of hidden and secret writing, the information are hidden in cover media (text, audio, image, video) in both spatial (LSB, VLSB) [1, 2, 3] and transform (DCT, DWT) domain [1]. In spatial domain LSB and VLSB [3, 4] are the most fashionable methods of Steganography [5]. In transform domain Steganography has been implemented using Discrete Wavelet Transform (DWT) [6] and Discrete Cosine Transform (DCT) e.g. adaptive DCT based mode 4 Steganography [8], pseudo-code algorithm [6] and compression [7].

The importance of steganography is increased day by day, and a lot of work is done to enhance this field. kaliappan gopalan proposed an audio steganography algorithm with an encryption key for the embedded secret audio file [8]. Mazdak Zamani, Hamed Taherdoost, Azizah A.Manaf, Rabiah B. Ahmad, and Akram M. Zeki work on higher LSB layer of audio file to increased its robustness [9]. Mohammad Pooyan, Ahmad Delforouzi use lifting wavelet transform and calculate the hearing threshold, and according to the threshold data bits are embedded in the least significant bits (LSB) of lifting coefficients [10]. Dimitriy E.Skopin, Ibrahim M.M. El-Emary, Rashad J.Rasras, Ruba S.Diab proposed an algorithm in audio steganography for hiding human speech signal using two methods spectrum shift and spectrum spreading [11]. Masahiro Wakivama, Yasunobu Hidaka, Koichi Nozaki work on wave file as an audio data, and proposed two kinds of new methods of extended low bits coding [12]. Haider Ismael Shahadi, Razali Jidin

proposed high capacity audio steganography algorithm based on the wavelet packet transform with adaptive hiding in least significant bits (LSB) [13]. Anupam Kumar Bairagi, Saikat Mondal, Amit Kumar Mondal proposed a method, the message bits are embedded into deeper layer in such a way to increase its robustness [14].

Audio Steganography is used to hide secret information in audio cover file the secret message is embedded by slightly altering the binary sequence of an audio file. Existing audio Steganography software can embed message in WAV, AU and even MP3 sound files, in both temporal domain and transform domain. In transform domain Steganography has been implemented using DWT and DCT etc.

In all previously implemented techniques a specific region/coefficients of the DCT are targeted and data/information are hidden in the least significant bits of specific DCT coefficients. This paper specifically deals with LSB Steganography using DCT. The main aim of this paper is to find and analyze the effect and contribution of each bit position on SNR and PSNR of Stego image in DCT domain and also to make a comparison with the effect and contribution of each bit position on SNR and PSNR of Stego image in spatial domain.

2. Quality Measuring Parameters

To analyze the effect of data hiding at different bit position in DCT coefficients SNR, PSNR and MSE are used for quality measuring. These parameters are used to compare the resulted Stego file with the original cover file and also with the resulted Stego files for data embedding at different positions.

These parameters used to measure the quality and error between cover audio and Stego audio are calculated using the following formulas [15, 16].

$$SNR = 10 * \log_{10} \left[\frac{\sum_{i=1}^R \sum_{j=1}^C [Cov(i, j)]^2}{\sum_{i=1}^R \sum_{j=1}^C [Cov(i, j) - Stego(I, j)]^2} \right] \quad (1)$$

$$MSE = \frac{1}{R * C} \sum_{i=1}^R \sum_{j=1}^C [Cov(i, j) - Stego(i, j)]^2 \quad (2)$$

$$PSNR = 10 * \log_{10} \left[\frac{2^{16} - 1}{MSE} \right] \quad (3)$$

3. Implementation

Hiding data in the least significant bits is a common practice in both spatial domain and transform domain. Here in this paper an analysis is made by hiding data at various bits position in the coefficients of discrete cosine transform of an audio cover file. A message audio file is hidden in a cover file in a cover audio file in .WAV format. The message and cover file are read at the sampling rate of 44100 samples/s and 351800 samples are captured almost equal to 3seconds play time.

To made analysis in transform domain DCT is applied on the cover audio file resulting in an array of $2 \times N$ size in Matlab. There is a problem of negative values occurring in Matlab and the DCT coefficients are scaled to avoid the negative values and make all coefficients positive. The coefficients are round off to fix the fraction part, by multiplying the DCT coefficient with a suitable number, without any loss. For example after DCT we get a value 0.3528 then we multiply it with 10^4 and get a whole number 3528. By direct rounding off diffidently some will be lost which avoided by the procedure adopted. Then information is hidden in the coefficients by substitution mechanism. The reverse of this process is applied back after hiding data at the specific bit position. The whole process is explained here in block diagram.

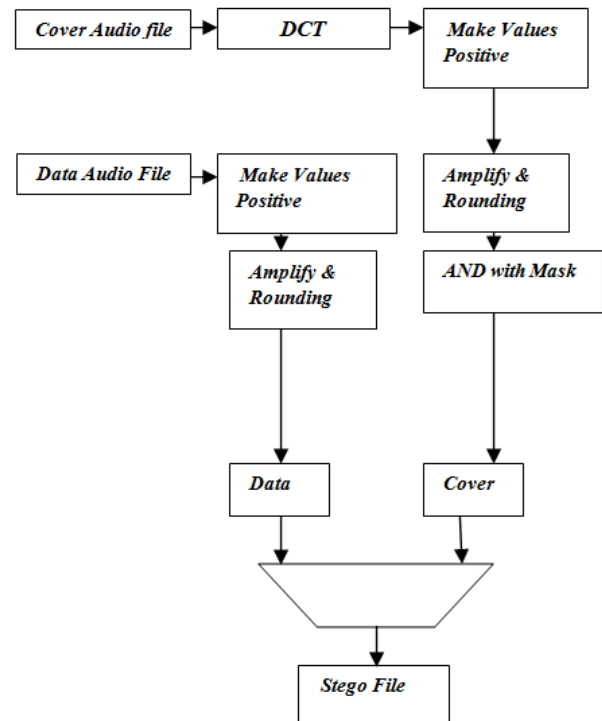


Figure 1: Block diagram Audio Steganography in DCT domain

The procedure is adopted for data hiding at various bit position and for each position a stego audio file is obtained and MSE, SNR and PSNR are calculated for each bit position.

4. Experimental Results

The procedure given in implementation section is implemented to get stego audio file for data hiding in 1st bit position, 2nd bit position, and 3rd bit position and so on. And for each bit position the MSE, SNR and PSNR are calculated using the expression given in the earlier section. Data hiding at various bit position create almost same, plus minus 1%, results of MSE, SNR and PSNR; quite different from image Steganography results [15]. The resulted values of MSE, PSNR, SNR, and Capacity are given in table 1. The estimated time elapsed in data hiding proves is also calculated for different bit position and a increasing trend has been observed as shown in table1.

Table1. MSE, PSNR, SNR, Capacity and Processing Time vs. Variable Bit Position

Bit position	MSE	PSNR (db)	SNR (db)	Capacity (%)	Processing Time (sec)
1 st	73.3746	89.7153	33.8824	5.8824	125.925
2 nd	73.3734	89.7154	33.6424	5.8824	128.963
3 rd	73.3722	89.7155	33.6423	5.8824	139.250
4 th	73.3657	89.7158	33.6419	5.8824	152.423
5 th	73.3650	89.7159	33.6419	5.8824	130.895
6 th	73.3554	89.7164	33.6413	5.8824	156.153
7 th	73.3363	89.7176	33.6402	5.8824	139.082
8 th	73.2320	89.7238	33.6340	5.8824	149.940

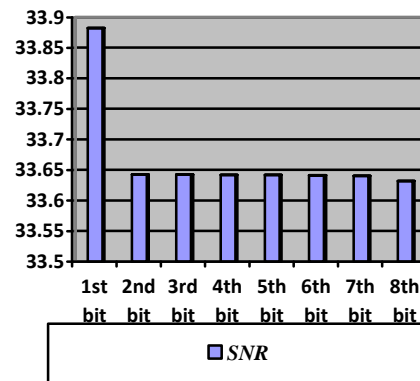


Figure 4: Analysis of signal to noise ration (SNR)

The resulted MSE, PSNR, SNR and time elapsed are also shown in figure 2, figure 3, figure 4 and figure 5 respectively for comparison.

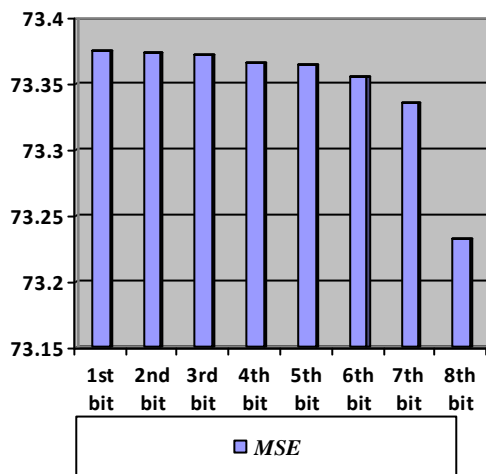


Figure 2: Analysis of mean square error (MSE)

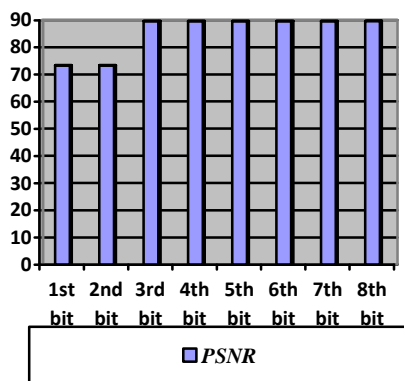


Figure 3: Analysis of peak signal to noise ratio (PSNR)

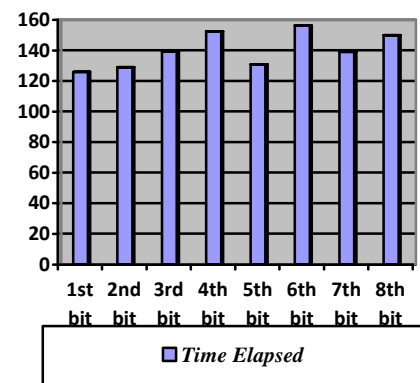


Figure 5: Analysis of Time elapsed

5. Conclusion

In this paper data is hidden at variable bit position in DCT coefficients of an audio cover file and the effect of each bit position is analyzed in term of MSE, SNR and PSNR. A very insignificant change with higher position occurs in the Stego file. Hiding data in DCT coefficients create a very significant noise in Stego file but the variation in SNR, PSNR and MSE is very insignificant as compared to that of Stegnography in DCT domain. For the same amount of noise as at 1st least significant bit position data can be hidden at 8th bit position.

References

- [1]. S. Dumitrescu, W.X.Wu and N. Memon, "On steganalysis of random LSB embedding in continuous-tone images", Proc. International Conference on Image Processing, Rochester 2002, NY, pp. 641-644.
- [2]. S.K. Moon and R.S. Kawitkar, "Data Security using Data Hiding", International Conference on Computational

Intelligence and Multimedia Applications 2007, pp.247-251.

- [3]. S. Khan, M. Haroon Yousaf and M. Jamal Akram, "Implementation of Variable Least Significant Bits Steganography using Decreasing Distance Decreasing Bits Algorithm", IJCSI Volume 8, Issue 6, November 2011.
- [4]. Sahib Khan and Mohammad Haroon Yousaf, "Implementation of VLSB Steganography Using Modular Distance Technique", Innovations and Advances in Computer, Information, Systems Sciences, and Engineering Lecture Notes in Electrical Engineering Volume 152, 2013, pp 511-525.
- [5]. Neeta Deshpande, Kamalapur Sneha, Daisy Jacobs, "Implementation of LSB Steganography and Its Evaluation for various Bits", Digital Information Management, 2006 1st International Conference on. 06/01/2007; DOI: 10.1109/ICDIM.2007.369349
- [6]. Sahib Khan and M. Haroon Yousaf, "Variable Least Significant Bits Steganography". (Accepted for publication in IJCSI Volume 8, Issue 5, September 2011.)
- [7]. K.B. Raja, C.R. Chowdary, Venugopal K R, L.M.Patnaik, "a secure image Steganography using LSB, DCT and compression technique on raw image" 0-7803-9588-3/05/\$20.00 ©2005 IEEE
- [8]. Muhammad Asad, Junaid Gilani, Adan Khalid, "An Enhanced Significant Bit Modification Technique for Audio Steganography", 2011 International Conference on Computer Networks and Information Technology (ICCNIT), pages 143-147
- [9]. kaliappan Gopalan, "A Unified Audio and image Steganography by Spectrum Modification", International Conference on Industrial Technology, 2009, page(s):1-5
- [10]. Hossein Malekmohamadi and shahrokh Ghaemmaghami "Reduced Complexity Enhancement of Steganalysis of LSB-matching Image Steganography" 2009 IEEE/ACS International conference on computer system and applications.
- [11]. N. Cvejic, T. Seppanen, "Increasing the capacity of LSB-based audio steganography." IEEE Workshop on Multimedia Signal Processing, pp. 336-338, 2002.
- [12]. K. Gopalan, "Audio steganography using bit modification," Proc. IEEE Int. conf. Acoustics, Speech, and Signal Processing, Vol. 2, pp. 421-424, April 2003.
- [13]. W.Bender, D. Gruhl, N. Morimoto and A. Lu, "Techniques for data hiding," IBM Systems Journal, Vol, 35, Nos. 3 & 4, pp.313-336, 1996.
- [14]. Pal S.K., Sexena P.K. and Mutto S.K. "The Future of Audio steganography". Pacific Rim Workshop on Digital Steganography, Japan, 2002.
- [15]. Sahib Khan, M. Nawaz Khan and Somia Iqbal, "Bit Position based Qualitative and Quantitative Analysis of DCT and Spatial Domain Steganography", IJCSI Volume 10, Issue 3, May 2013
- [16]. <http://www.mathworks.com>



Engr. Sahib Khan got M.Sc. in Telecommunication Engineering from University of Engineering and Technology, Taxila in the year 2012 and B.Sc. in Telecommunication Engineering from University of Engineering and Technology, Peshawar in the year 2008. He currently holds the position of Lecturer in the Department of Electrical Engineering at University of Engineering and Technology, Peshawar. His research interests include

Image and Signals Processing, Information Security and Antenna designing. He, along with his students, is developing and implementing Steganographic techniques and algorithms, esp. Image and Audio Steganography, in Spatial and Transform domains.

The Effect of Various Number of Least Significant Bits substitution in Audio using Discrete Cosine Transform

Sahib Khan¹, Umar Said², Ejaz Ahmad³, Fazlullah⁴ and Mukhtair Ali⁵

¹⁻⁵ Dept. of Electrical Engineering, University of Engineering and Technology, Peshawar, Pakistan

Abstract

A secure exchange of information is the need of today's digital world, achieved by either changing the shape of information according to a predefined key using cryptography or concealing the information in a cover media, in spatial or transform domain, using Stegnography. The hiding data in various least significant bits of DCT coefficients of an audio cover file is the subject of this work. This paper statistically analyze the effect of different number of least significant bits (LSB) on Stego audio in terms of signal to noise ratio (SNR), peak signal to noise ratio (PSNR), mean square error (MSE) and data hiding capacity. The hiding capacity increases significantly with the increase in the number of least significant bit substituted with information bits and results in a very insignificant variation in SNR, PSNR and MSE. The statistical analysis shows that audio Stegnography in DCT domain is subjected to more distortion at lower hiding capacity and is subject to almost the same distortion for large data hiding capacity.

Keywords: *Audio Stegnography, Discrete Cosine Transform, Watermarking, Information Security, Steganalysis*

1. Introduction

The infrastructure for distribution of digital media has grown rapidly. This provides an excellent opportunity. There for we have so many methods for data hiding i.e. Stegnography, cryptography, watermarking. Stegnography is a technique used for hiding information in a cover media such as audio, video, text and image. In both temporal (LSB, VLSB) [1, 2, 3] and transform domain (DCT, DWT) [1]. The common methods that are used for data hiding in temporal domain are LSB and VLSB [3, 4]. In transform domain, we can hide data by using Discrete Wavelet Transform (DWT) and Discrete Cosine Transform [5, 6]. That is DCT based on mode 4 Stegnography [8], pseudo code algorithm and compression [6, 7].

The importance of Stegnography increases day by day, so the work and research is also increase in this field. There is lots of research work related audio Stegnography going on by various institutes and individuals. Some papers about dual Stegnography are also proposed. That is using cryptography for providing additional security. All the authors tried to get high capacity and high security.

In kaliappan gopalan proposed a steganalysis in audio file with an encryption key for the embedded secret audio file [9]. In M Asad, J Gilani and A Khalid proposed a audio steganography with an encrypted audio file using advanced Encryption Standard(AES) [10]. In Mazadak Zaman et. al proposed a genetic algorithm for the embedding secret audio files for achieving higher robustness and capacity [11]. In kaliappan gopalan embed the information of discrete message in spectral domain of a cover audio of image file [12]. In K.B.Raja et. al. proposed a work on image steganography where a LSB embedding is used and then DCT is performed followed by a compression technique to provide high security in the hidden data [13]. In Hossein Malekmohamadi and Shakrokh Ghaemmaghani proposed an enhancement in image steganalysis of LSB matching by reducing the complexity using gober filter coefficients [14]. In R Balagi and G Naveen extended their work towards video steganography by embedding the secret information in some particular frames [15]. In Andrew D. Ker derived a mathematical analysis for the steganalysis in last two LSB bits [16]. S. Khan, M. Haroon Yousaf and M.Jamal Akram proposed an algorithm for Implementation of variable Least Significant Bit steganography [3]. A very related work in DCT domain has recently been published by Sahib Khan et. al using image as cover file [17]

The main task of this paper is to hide data in various number of least significant bits of DCT coefficients' of an audio cover file and analyze the effect of each combination in terms of SNR, PSNR, MSE and data hiding capacity.

is a technique of hidden and secrete writing, the information are hidden in cover media (text, audio, image, video) in both spatial (LSB, VLSB) [1, 2, 3] and transform (DCT, DWT) domain [1]. In spatial domain LSB and VLSB [3, 4] are the most fashionable methods of Stegnography [5]. In transform domain Stegnography has been implemented using Discrete Wavelet Transform (DWT) [6] and Discrete Cosine Transform (DCT) e.g. adaptive DCT based mode 4 Stegnography [8], pseudo-code algorithm [6] and compression [7].

The importance of steganography is increased day by day, and a lot of work is done to enhance this field. kaliappan gopalan proposed an audio steganography algorithm with an encryption key for the embedded secret

audio file [8]. Mazdak Zamani, Hamed Taherdoost, Azizah A.Manaf, Rabiah B. Ahmad, and Akram M. Zeki work on higher LSB layer of audio file to increased its robustness [9]. Mohammad Pooyan, Ahmad Delforouzi use lifting wavelet transform and calculate the hearing threshold, and according to the threshold data bits are embedded in the least significant bits (LSB) of lifting coefficients [10]. Dimitriy E.Skopin, Ibrahim M.M. El-Emary, Rashad J.Rasras, Ruba S.Diab proposed an algorithm in audio steganography for hiding human speech signal using two methods spectrum shift and spectrum spreading [11]. Masahiro Wakivama, Yasunobu Hidaka, Koichi Nozaki work on wave file as an audio data, and proposed two kinds of new methods of extended low bits coding [12]. Haider Ismael Shahadi, Razali Jidin proposed high capacity audio steganography algorithm based on the wavelet packet transform with adaptive hiding in least significant bits(LSB) [13]. Anupam Kumar Bairagi, Saikat Mondal, Amit Kumar Mondal proposed a method, the message bits are embedded into deeper layer in such a way to increase its robustness [14].

Audio Stegnography is used to hide secret information in audio cover file the secret message is embedded by slightly altering the binary sequence of an audio file. Existing audio Stegnography software can embed message in WAV, AU and even MP3 sound files, in both temporal domain and transform domain. In transform domain Stegnography has been implemented using DWT and DCT etc.

In all previously implemented techniques a specific region/coefficients of the DCT are targeted and data/information are hidden in the least significant bits of specific DCT coefficients. This paper specifically deals with LSB Stegnography using DCT. The main aim of this paper is to find and analyze the effect and contribution of each bit position on SNR and PSNR of Stego image in DCT domain and also to make a comparison with the effect and contribution of each bit position on SNR and PSNR of Stego image in spatial domain.

2. Analysis Parameters

To analyze the effect of different number of least significant bits in DCT coefficients of an audio cover file the SNR, PSNR and MSE are calculated for each combination of least significant bits and the data hiding capacity for each combination is also calculated.

2.1. Hiding Capacity

The information hiding capacity for each combination of least significant bits used is calculated using the expression given.

$$C = \frac{\sum_{i=1}^N (Cf_i \times Bi)}{N \times 8} \times 100$$

Where

N: Size of Cover file

Bi: The number of bits hidden in a Coefficient

Cfi: The ith Coefficient

2.2. MSE, SNR and PSNR

The quality of Stego Audio file is analysed quantitatively by calculating mean square error (MSE), signal to noise ratio (SNR) and peak signal to noise ratio (PSNR). The MSE, SNR and PSNR are calculated using expressions given below [18, 19, 20].

$$MSE = \frac{1}{R * C} \sum_{i=1}^R \sum_{j=1}^C [Cov(i, j) - Stego(i, j)]^2$$

$$SNR = 10 * \log_{10} \left[\frac{\sum_{i=1}^R \sum_{j=1}^C [Cov(i, j)]^2}{\sum_{i=1}^R \sum_{j=1}^C [Cov(i, j) - Stego(I, j)]^2} \right]$$

$$PSNR = 10 * \log_{10} \left[\frac{2^{16} - 1}{MSE} \right]$$

3. Implementation

Hiding data in the least significant bits is a common practice in both spatial domain and transform domain. This work presented in this paper analyze the effect of various number of least significant bits substituted in the DCT coefficients of an audio file. An audio message file is hidden in the least significant bits of the DCT coefficients of cover audio file. The message and cover file are read at the sampling rate of 44100 samples/s and 351800 samples are captured almost equal to 3seconds play time in Matlab. After reading/recording the cover file discrete cosine transform (DCT) is applied on the cover audio file resulting in a group of DCT coefficients. There is a problem of negative values occurring in Matlab and the DCT coefficients are scaled to avoid the negative values and make all coefficients positive. The coefficients are round off to fix the fraction part, by multiplying the DCT coefficient with a suitable number, without any loss. For

example after DCT we get a value 0.3528 then we multiply it with 10^4 and get a whole number 3528. By direct rounding off differently some will be lost which avoided by the procedure adopted. Then information is hidden in the specific least significant bits. The whole process is explained in the block diagram in figure 1.

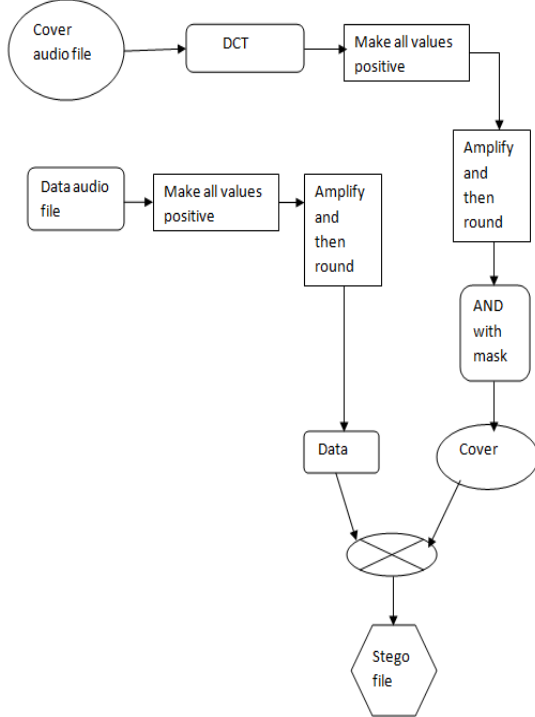


Figure.1: Block diagram of Audio Stegnography in DCT

To make the analysis for various numbers of least significant bits, data is hidden in 1LSB, 2LSBs, 3LSBs, and 4LSBs and so on. For each combination of least significant bits SNR, PSNR, MSE and hiding capacity are calculated.

4. Experimental Results

To analyze the effect of various numbers of least significant bits on SNR, PSNR, MSE and hiding capacity, different numbers of least significant bits of the DCT coefficients of audio file are modified according to the message/information. For each combination of least significant bits the quality measuring parameters and hiding capacity is calculated. It can be clearly observed from the experimental results that hiding capacity significantly increases with the increase in number of least significant bits to substituted in cover file as shown in table 1 while the SNR, PSNR and MSE doesn't changes significantly with the increase in number of bits this trend is quite opposite than that of using image as cover. The values of SNR, PSNR and MSE are also given in table 1.

The time elapsed is also calculated for each combination of least significant bits and is given in table 1. For comparison the SNR, PSNR, MSE, Hiding capacity and estimated time (elapsed time) for each combination of LSBs are shown graphically in figure 2, figure 3, figure 4, figure 5 and figure 6 respectively.

Table1. MSE, PSNR, SNR, Capacity and Processing Time vs LSBs

No of LSBs	SNR (db)	PSNR (db)	MSE (db)	Elapsed time in sec	Capacity in %
1LSB	32.0691	83.7869	71.8334	33.880674	5.8824
2LSBs	32.0690	83.7870	71.8320	35.138149	11.7647
3LSBs	32.0687	83.7872	71.8280	34.358654	17.6471
4LSBs	32.0684	83.7875	71.8228	34.249364	23.5294
5LSBs	32.0678	83.7881	71.8125	34.908435	29.4118
6LSBs	32.0666	83.7894	71.7920	34.769906	35.2941
7LSBs	32.0641	83.7918	71.7511	35.124846	41.1765
8LSBs	32.0591	83.7968	71.6688	35.267543	47.0588

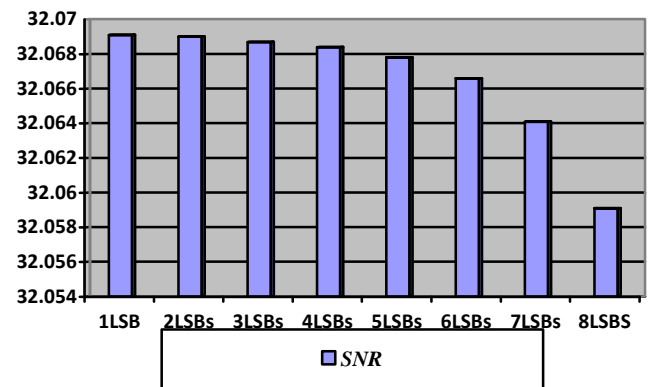


Figure 2: SNR vs. No. of Least Significant Bits

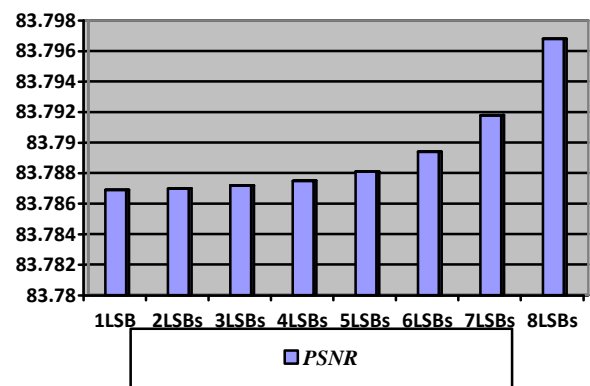


Figure 3: PSNR vs. No. of Least Significant Bits

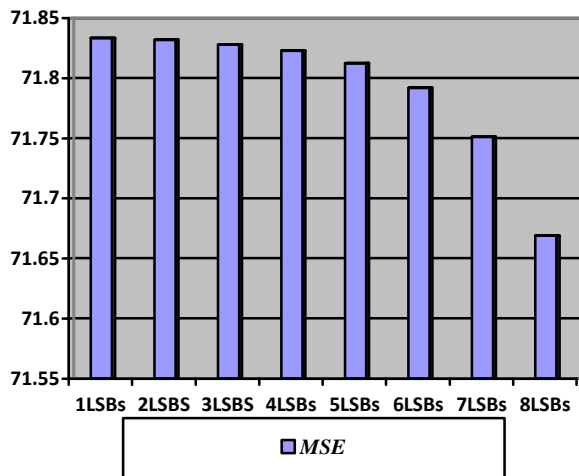


Figure 4: MSE vs. No. of Least Significant Bits

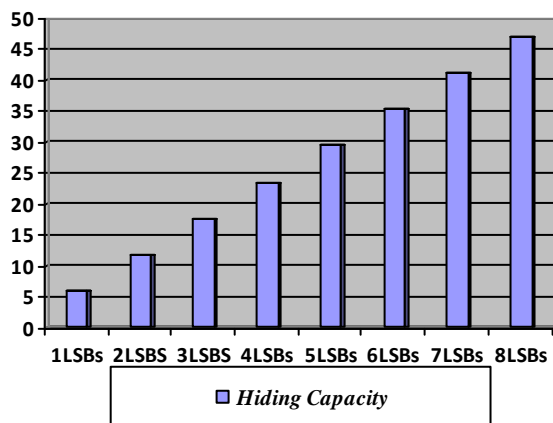


Figure 5: Hiding Capacity vs. No. of Least Significant Bits

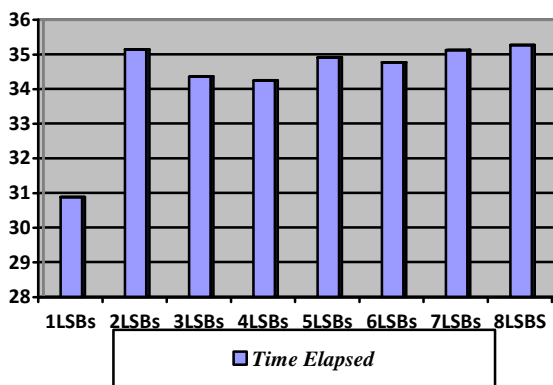


Figure 6: Time Elapsed vs. No. of Least Significant Bits

5. Conclusion

Hiding data in different number of LSBs of DCT coefficients of an audio cover file show a very significant distortion in the cover file for less hiding capacity while almost the same distortion is created for large data hiding capacity of 47 %. The results of SNR, PSNR and MSE show that Audio Steganography in DCT domain is very suitable for large data hiding instead of less data hiding

References

- [1]. S. Dumitrescu, W.X.Wu and N. Memon, "On steganalysis of random LSB embedding in continuous-tone images", Proc. International Conference on Image Processing, Rochester 2002, NY, pp. 641-644.
- [2]. S.K. Moon and R.S. Kawitkar, "Data Security using Data Hiding", International Conference on Computational Intelligence and Multimedia Applications 2007, pp.247-251.
- [3]. S. Khan, M. Haroon Yousaf and M. Jamal Akram, "Implementation of Variable Least Significant Bits Steganography using Decreasing Distance Decreasing Bits Algorithm", IJCSI Volume 8, Issue 6, November 2011.
- [4]. Sahib Khan and Mohammad Haroon Yousaf, "Implementation of VLSB Steganography Using Modular Distance Technique", Innovations and Advances in Computer, Information, Systems Sciences, and Engineering Lecture Notes in Electrical Engineering Volume 152, 2013, pp 511-525.
- [5]. Neeta Deshpande, Kamalapur Sneha, Daisy Jacobs, "Implementation of LSB Steganography and Its Evaluation for various Bits", Digital Information Management, 2006 1st International Conference on. 06/01/2007; DOI: 10.1109/ICDIM.2007.369349
- [6]. Sahib Khan and M. Haroon Yousaf, "Variable Least Significant Bits Steganography". (Accepted for publication in IJCSI Volume 8, Issue 5, September 2011.)
- [7]. K.B. Raja, C.R. Chowdary, Venugopal K R, L.M.Patnaik, "a secure image Steganography using LSB, DCT and compression technique on raw image" 0-7803-9588-3/05/\$20.00 ©2005 IEEE
- [8]. Muhammad Asad, Junaid Gilani, Adan Khalid, "An Enhanced Significant Bit Modification Technique for Audio Steganography", 2011 International Conference on Computer Networks and Information Technology (ICCNIT), pages 143-147
- [9]. Gopalan., "Audio steganography using bit modification", 2003 IEEE international conference on Acoustic, speech and signal processing page(s):II -421 -4 vol.2.
- [10]. Muhammad Asad, Junaid Gilani, Adan Khalid "An Enhanced Significant Bit Modification Technique for Audio Steganography", 2011 international conference on computer Networks and Information Technology (ICCNIT), pages 143-147
- [11]. Zamani M., Manaf, A, Ahmad, R.B Jaryani, F., Taherdoost H., Zeki, AM., "A secure audio steganography approach",

- International Conference for internet Technology and Secured Transactions 2009, page(s):1-6.
- [12].Kaliappan Gopalan, "A Unified Audio and image Steganography by Spectrum Modification", International Conference on Industrial Technology, 2009, page(s):1-5
- [13].K B Raja , Chowdary C R, Venugopal K R, Patnaik L M "A Secure Image Steganography using LSB DCT and compression Techniques on Raw Images" 2005 IEEE International conference on session B-image signal processing.
- [14].Hossein Malekmohamadi and Shahrokh Ghaemmaghami "Reduced Complexity Enhancement Of Steganalysis of LSB-matching Image Steganography" 2009 IEEE/ACS International conference on computer system and applications.
- [15].R Balagi, Naveen G"Secure Data Transmission Using Video Steganography", 2011 IEEE International conference on electro/information technology (EIT).
- [16].Andrew D.Ker,"Steganalysis of Embedding in Two Least-Significant bits" 2007 IEEE transactions on information forensics and security, vol.2 no.1, march 2007.
- [17].Dmitriy E.Skopin, Ibrahim M. M. EL-Emary, Rashad J.Rasras, Ruba S.Diab, "Advanced Algorithm in Audio Steganography for Human Speech Signal".
- [18].<http://www.mathworks.com>
- [19].R.C. Gonzalez and R.E. Woods, 2008. Digital Image Processing. Third Edition, Prentice Hall2008. ISBN: 013168728.
- [20].Sahib Khan, M. Nawaz Khan and Somia Iqbal, "Bit Position based Qualitative and Quantitative Analysis of DCT and Spatial Domain Steganography", IJCSI Volume 10, Issue 3, May 2013



Engr. Sahib Khan got M.Sc. in Telecommunication Engineering from University of Engineering and Technology, Taxila in the year 2012 and B.Sc. in Telecommunication Engineering from University of Engineering and Technology, Peshawar in the year 2008. He currently holds the position of Lecturer in the Department of Electrical Engineering at University of Engineering and Technology, Peshawar. His research interests include Image and Signals Processing, Information Security and Antenna designing. He, along with his students, is developing and implementing Stegnographic techniques and algorithms, esp. Image and Audio Stegnography, in Spatial and Transform domains.

The True Powers of Multi-core Smartphones

Dhuha Basheer Abdullah¹, Mohammed M. Al-Hafidh²

¹Computer Science Dept., Mosul University, Computer Sciences and Mathematics College
Mosul, Iraq

²Computer Science Dept., Mosul University, Computer Sciences and Mathematics College
Mosul, Iraq

Abstract

Just as the multi-core processors took all the attention in general computing field in the past decade, so too multi-core smartphones are taking all the attention today. Relying on the principles of tasks and data parallelism. We propose in this paper a parallel programming approach on quad-core smartphones to do a big matrices multiplication, and show how to increase the utilization of processors to achieve improvement on the system's runtime.

Keywords: Smartphone, Parallel processing, Multi-core processor, Matrices multiplication.

1. Introduction

Mobility is a key term in the world of today. No matter where you are or what you are doing you are surrounded by a world of mobile devices. No longer are we confined to work at a desktop terminal, we can now work and communicate from virtually any location. This new mode of interaction has been made possible by the advances in the world of miniaturisation. We can now work and communicate by using a myriad of devices such as: Laptops, Ultra Mobile PC's, PDA's and mobile phones. Even though we may be surrounded by these devices, the question must be raised: are we using them to their fullest potential? One possible solution to this lies within the area of parallel processing [10], without realizing the true potential of embedded multi-core architecture, we are not making full use of this now technology.

Multi-core architectures for personal computers is an important field in our everyday computing, many frameworks and APIs focusing on parallel programming have been proposed for multi-core processors, that achieve speedup and maximum processor utilization. However, very little researches have been proposed in the area of multi-core architecture for smartphones primarily because this is a relatively new concept and not many end-products have embraced such architecture. The recent release of multi-core smartphones optimized for both performance

and power consumption, such as Samsung Galaxy SII and SIII [8], has revolutionized mobile computing and opened up the door to new research paradigms, especially for real-time processing.

Now we can consider complex algorithms for potential implementation that previously regarded as impractical for deployment on smartphones platforms. For instance, performing certain computation on big data matrices on smartphone was very difficult in the past, due in part to its memory constraints but primarily to the processing power of the smartphone. It is desirable to be able to use complex algorithms whenever possible because they generally yield more accurate results. Fortunately, the recent release of multi-core smartphones has empowered us to do exactly that, as true parallelism can now be achieved [1].

2. Contributions

In this paper, we consider the problem of big matrices multiplication, relying on the principles of task and data parallelism, we propose in this paper a parallel programming approach on quad-core mobile devices to do big matrices multiplication that has the following properties:

1. Show how to increase the utilization of processors to achieve improvement on the system's runtime.
2. Show the full capabilities of multi-core processor utilization.

3. Related work

On the parallel programming front, making a task parallelizable and run on multiple cores can be a grueling process. Challenging issues include thread synchronization, data race, and starvation.

Many attempts have been made by researchers and programmers alike to design a high-level framework that provides an abstraction layer for programmers to use. Such framework allows the programmers to fully focus on application development without unnecessary worry about parallel programming. The ParLab at Berkeley, UPCRC at

Illinois, and the Pervasive Parallel Laboratory at Stanford propose a two-layer framework, consisting of the productivity layer where domain experts, assumed to have limited experience with parallel programming, can focus on application development, and the efficiency layer where computer scientists with strong background in parallel programming can focus on improving the efficiency of the application [5].

Similar to the aforementioned framework, programming models such as algorithmic skeletons have also been proposed, aiming to benefit from multi-core architectures while decoupling the hassle of thread management from common programming. Skandium and Calcium [4] provide high-level parallel programming libraries based on the thread pool and ExecutorService frameworks in JAVA. Users only need to provide a threshold for threads and a set of initial parameters.

Daniel C. Doolan, and Laurence T. Yang in year 2006. They considered the problem of matrix multiplication to show and demonstrates that mobile devices are capable of parallel computation using Mobile Message Passing Interface (MMPI). MMPI allows parallel programming of mobile devices over a Bluetooth network [2].

Panya Chanawangsa, and Chang Wen Chen in year 2012. They demonstrate how proper utilization of a dual-core mobile processor can achieve tremendous speedup in mobile application [1].

4. Development Platform

The system was developed on Android 4.1.2 (Jelly Bean). Released in September 2012. The phone's most outstanding feature is its processor – a superscalar quad-core 1.4 GHz Arm Cortex-A9 with 2 GB RAM [8]. Optimized for high performance and low power consumption, the Galaxy SIII is indeed an ideal platform for this system.

5. Matrix Multiplication

Matrix multiplication is a time consuming operation because for an $n \times n$ matrix the best possible time complexity is $O(n^2)$ it cannot be less than this as all n^2 cells must be visited. Assume two matrices are to be multiplied, if A is an $n \times m$ matrix and B is an $m \times p$ matrix, the result AB of their multiplication is C $n \times p$ matrix defined only if the number of columns m in A is equal to the number of rows m in B. When multiplying matrices, the elements of the rows in the first matrix are multiplied with corresponding columns in the second matrix. One may compute each entry in the third matrix one at a time [2][3]. For two matrices

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1p} \\ B_{21} & B_{22} & \cdots & B_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ B_{m1} & B_{m2} & \cdots & B_{mp} \end{pmatrix}$$

(where necessarily the number of columns in **A** equals the number of rows in **B** equals m) the matrix product **AB** is defined by

$$AB = \begin{pmatrix} (AB)_{11} & (AB)_{12} & \cdots & (AB)_{1p} \\ (AB)_{21} & (AB)_{22} & \cdots & (AB)_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ (AB)_{n1} & (AB)_{n2} & \cdots & (AB)_{np} \end{pmatrix}$$

where **AB** has entries defined by

$$(AB)_{ij} = \sum_{k=1}^m A_{ik}B_{kj}. \tag{1}$$

The implementation of matrix multiplication in application code is generally a question of three loops as shown in

```
private void mult(int a[][],int b[][],
int c[][])
{for (int i = 0; i < a.length; i++)
{for (int j = 0; j < b[0].length; j++)
{for (int k = 0; k < a[0].length; k++)
{c[i][j] += a[i][k] * b[k][j];}}}}
```

Figure (1).

Fig. 1 Method to calculate product of two matrices.

6. Parallel Processing

Parallel processing involves multiple processes which are active simultaneously and solving a given problem, generally on multiple processors. We are underplaying the role of physical multiprocessing here, because the study of parallel processing does not require multiple processors. Most of OS today provide multiprocessing / tasking, which can be exploited to study the issues and programming aspects of pp. But, we must have multiple processes (or rather, independent execution units) which are simultaneously active. These units take the role of different processing units (or processors). The critical aspect here is “solving a given problem”. The processes must all be concerned with the solution of one single problem. In other words, there must be interaction among the units. For example, one unite compiling file1 and another compiling file2, will not be considered as parallel processing, because these two are absolutely independent

tasks according to the definition of C language. But, if one unit is compiling a statement from file1 and the other is compiling the next statement from the same file, they need to share variable declaration and scoping information, properly between them. Hence they are solving a single problem and not two separate problems. We can, therefore, consider that as parallel processing [9].

7. Multithreading and Processor Utilization

Since a mobile phone is considered a general-purpose device, application-level parallelism is the best we can achieve. Without explicitly using multiple threads, speedup from a multi-core architecture will not be obvious. In this paper, we propose a general guideline for breaking down a global task into multiple subtasks and later demonstrate how to apply this idea on a matrix multiplication task.

The first step towards parallelizing a task is to determine the optimal number of threads to use. Limiting thread contention is crucial for application speedup. Spawning too many threads than necessary not only disrupts other applications, but may also result in a longer execution time of the application due to the overhead associated with context-switching. A processor core can handle only one thread at a time. For efficiency purposes, a simple rule is to spawn as many threads as the number of cores available, thereby delegating one thread to each core and eliminating the need for time-slicing. A simulation was conducted by spawning different numbers of threads to execute certain tasks. A dramatic improvement in execution time can be seen when we increase the number of worker threads from 1 to 4. However, since there are only four available cores, increasing the number of threads does not enhance but aggravates the performance, resulting in a slightly longer execution time.

8. Matrix slicing

The next step is to determine if data parallelism is possible and appropriate. For this work, matrix multiplication, data comes in the form of matrices, or on the lower level 2-dimensional arrays. In many cases, they can be split up into smaller independent chunks and processed concurrently, reducing the execution time while producing the same output as when processed sequentially. For instance, to multiply two matrices A and B, we can split them up into smaller sub-matrices (matrix slicing) and multiply each sub-matrix from matrix A with the corresponding sub-matrix from matrix B concurrently. However, if the size of data to be processed is not significantly large, employing data parallelism will not yield much speedup as a result of thread overhead. In the case of this application, since the matrices are large (2000×2000), data parallelism is well worth a try. Generally speaking, given k processor cores, we should divide the input data of size n into n/k smaller

chunks and distribute them across k cores with each core running a single thread. Figure (2).

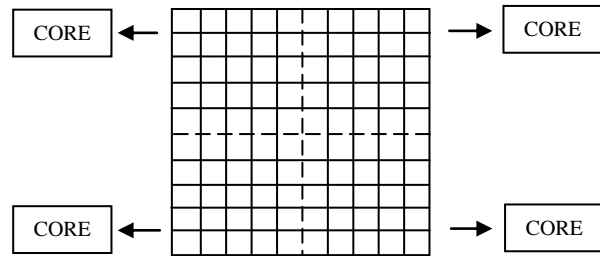


Fig. 2 Matrix slicing

9. Speedup

Speedup is the expected performance benefit from running an application on a multi-core versus a single-core machine. When speedup is measured, single-core machine performance is the baseline. For example, assume that the duration of an application on a single-core machine is six hours. The duration is reduced to three hours when the application runs on a quad machine. The speedup is 2-(6/3)- in other words, the application is twice as fast [7]. Speedup can be found by using the formula below

$$\text{Total Speedup} = T_s / T_p \quad (2)$$

T_s : is the runtime without parallelism.

T_p : is the runtime with parallelism.

10. Putting It All Together: The Proposed System In Action

For the parallel computation we consider that the matrices are square. The parallel computation requires several steps that are not required in the sequential version of the application (Figure 3). One of the first main differences is that we need to determine the number of threads suitable for the phone hardware. Since Samsung Galaxy SIII has four cores, four threads are ideal for execution the task in hand. So that the matrices must be split in to four same size square blocks called sub-matrix, then send each sub-matrix to one of the four cores we have. Once this operation has been completed each core can compute its own section of the matrix. The final stage is to gather all the results back in to one matrix. The result of the computation can then be available for the user.

To gain full control of task management, we also make use of the FutureTask class, allowing us to track the progress of the submitted tasks and block until all of them have been completed. The four tasks are eventually submitted to

an `ExecutorService`, which takes care of thread pool creation and assigns a submitted task to an available thread. More importantly, by using the `ExecutorService`, memory consistency is guaranteed, thus eliminating the trouble of thread synchronization. Figure (4).

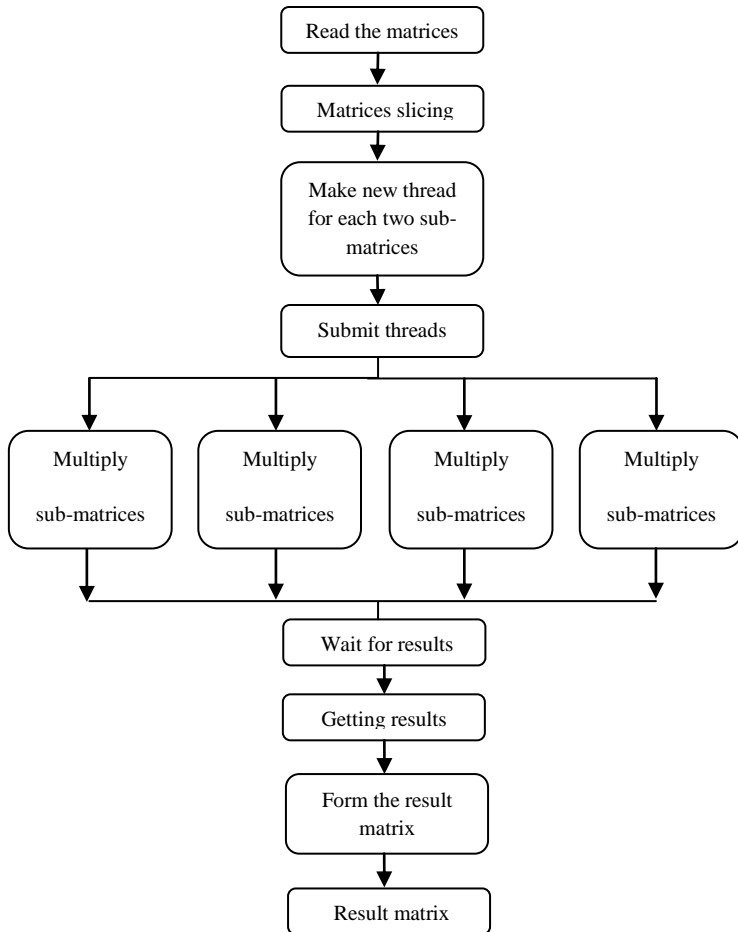


Fig. 3 Overall functionality of the system.

11. Experimental Results

The system tested on different size matrices, The results are shown in (Table (1)) below:

Table 1: Experimental results

Matrix size	Time in sequential	Time in parallel
500×500	7,657ms	231ms
1000×1000	99,911ms	1336ms
2000×2000	1,116,199ms	9092ms

```

Final ExecutorService es;
int nsize=size/2;
int[][] a11=new int[nsize][nsize];
int[][] a12=new int[nsize][nsize];
int[][] a21=new int[nsize][nsize];
int[][] a22=new int[nsize][nsize];
for(int i =0;i<nsize;i++)
for(int j=0;j<nsize;j++)
    {a11[i][j]=a[i][j];
    a12[i][j]=a[i+nsize][j];
    a21[i][j]=a[i][j+nsize];
    a22[i][j]=a[i+nsize][j+nsize];}
Future f1; Future f2; Future f3; Future f4;
f1=es.submit(new Callable<Object>(){
@Override
public Object call() throws Exception
{mult1(a11);
return true; }});
f2=es.submit(new Callable<Object>(){
@Override
public Object call() throws Exception
{mult2(a12);
return true; }});
f3=es.submit(new Callable<Object>(){
@Override
public Object call() throws Exception
{mult3(a21);
return true; }});
f4=es.submit(new Callable<Object>(){
@Override
public Object call() throws Exception
{mult4(a22);
return true; }});
try {
f1.get();f2.get();f3.get();f4.get();
Toast.makeText(getBaseContext(),
"Getting results",Toast.LENGTH_SHORT).show();
} catch (InterruptedException e) {
e.printStackTrace();}
    
```

Fig. 4 Code snippet illustrating matrix slicing, task submission and use of `ExecutorService`.

Figure (5) below shows runtime comparison between sequential and parallel.

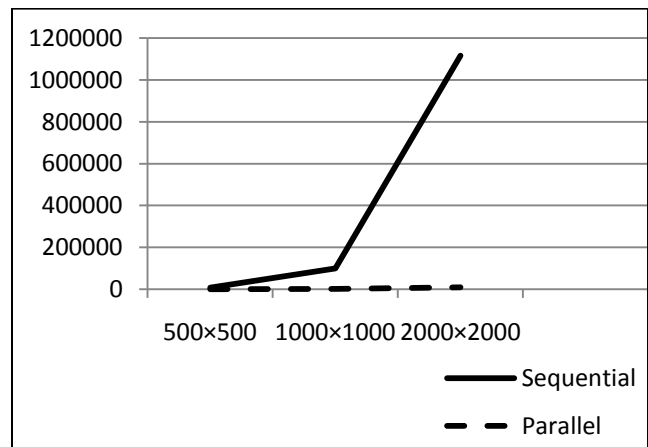


Fig 5. Runtime comparison between sequential and parallel.

Figure (6) below shows expected performance comparison.

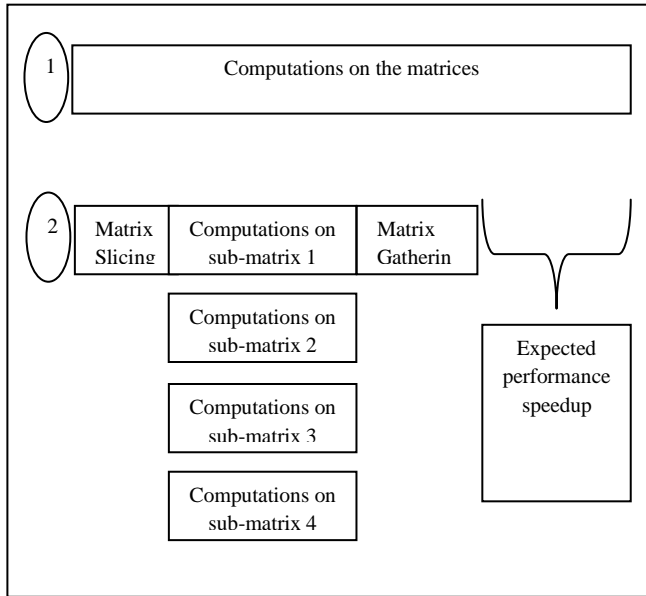


Fig. 6 Expected performance comparison : (1) Matrix multiplication without using the proposed system vs. (2) Matrix multiplication using the proposed system.

From the time results in (Table (1)), we can see that the system achieved a great speedup as shown in (Table (2)) below:

Table 2: Speedup results

Matrix size	Speed up
500×500	33.147
1000×1000	74.783
2000×2000	122.767

Figure (7) below shows speedup results.

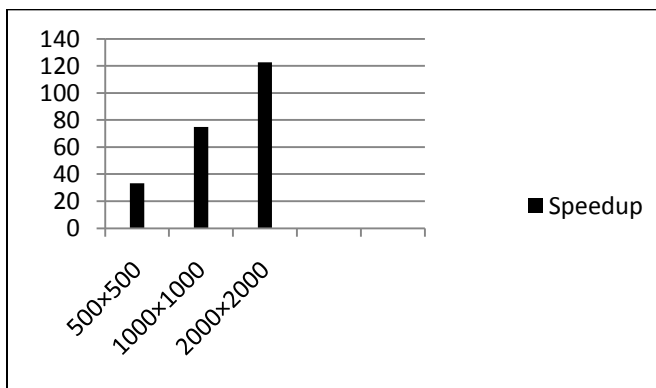


Fig 7. Speedup results

12. Conclusions

In this paper, we have demonstrated how to achieve speedup in a big matrices multiplication system written entirely in JAVA by using the proposed parallel programming approach, based on the idea of task and data parallelism. Running on a quad-core Samsung Galaxy SIII. The proposed system shows significant reduction in the overall processing time and great speedup.

13. References

- [1] Panya Chanawangsa, Chang Wen Chen, “A New Smartphone Lane Detection System: Realizing TruePotential of Multi-core Mobile Devices”, MoVid’12, 2012, pp.19-24.
- [2] Laurence T. Yang, Daniel C. Doolan, “Mobile Parallel computing”, Proceedings of The Fifth International Symposium on Parallel and Distributed Computing, IEEE International,2006.
- [3] Gene H. Golub, Charles F. Van Loan, Matrix Computations, The Johns Hopkins University Press,2013.
- [4] Tsogkas Panagiotis, “Evaluating Skandium’s Divide-and-Conquer Skeleton”, Master Thesis, School of Information, University of Edinburgh, 2010.
- [5] Catanzaro, B., et al., “Ubiquitous Parallel Computing from Berkeley, Illinois, and Stanford”, IEEE Computer Society, 2010, pp. 41-55.
- [6] Herve Guihot, Pro Android Apps Performance Optimization, Apress, 2012.
- [7] Donis Marshall, Parallel Programming with Microsoft Visual Studio, Microsoft Corporation by: O’Reilly Media, 2011.
- [8] Samsung I9305 Galaxy S III Full Specifications, http://www.gsmarena.com/samsung_i9305_galaxy_s_iii-5001.php
- [9] M. Sasikumar, Dinesh Shikhare, P. Ravi Prakash, Introduction To Parallel Processing, Prentice-Hall of India Private Limited, 2006.
- [10] Daniel C Doolan, Sabin Tabirca, Laurence T Yang,” MMPI a Message Passing Interface for the Mobile Environment”, Proceedings of MoMM2008, Linz, Austria, 2008.

Dhuha Albazaz is the head of Computer Sciences Department, College of Computers and Mathematics, University of Mosul. She received her PhD degree in computer sciences in 2004 in the speciality of computer architecture and operating system. She supervised many Master degree students in operating system, computer architecture, dataflow machines, mobile computing, real time, and distributed databases. She supervised three PhD students in FPGA field, distributed real time systems, and Linux clustering. She also leads and teaches modules at both BSc, MSc, and PhD levels in computer science. Also, she teaches many subjects for PhD and master students.

Mohammed M. Al-Hafidh is a master student in Computer Sciences Department, College of Computers and Mathematics, University of Mosul. He interest with networks, Databases, and operating system subjects

An Expert System for Diabetic Microvascular Complication Diagnosis

Truc Ly Thi Nguyen¹, Nhon Van Do²

¹ Department of Computer Science, Industrial University of Ho Chi Minh City,
Ho Chi Minh City, Vietnam

² Department of Computer Science, University of Information Technology,
Ho Chi Minh City, Vietnam

Abstract

This work presents the design of an expert system that aims to provide the doctors with limited practical experience and the diabetic patients more background for suitable diagnosis of diabetic microvascular complications. In modern society, more and more people get diabetes. This disease affects almost every organ in the body like heart, eyes, kidney, skin, nerves, blood vasculum and foot etc. They leave the very serious consequences for the patient. Early diagnosis of complications of diabetes can get tremendous value to patients and society. Hence an expert system for diabetes microvascular complication diagnosis can be used to assist for the doctors with limited clinical experience in situations without consultation directly to a specialist. In constructing an expert system, the knowledge representation plays a very important key. Therefore in this paper we also propose a new model of knowledge representation used to represent the complicated knowledge in medical domain and construct the Expert System for Diabetic Microvascular Complication Diagnosis (ESDMCD).

Keywords: *Knowledge base system, Knowledge representation, intelligent system, Expert system, Medical diagnosis.*

1. Introduction

According to specialist Nguyen Thy Khue - Dean of the Endocrinology Department of Medical University, said in the document [8], the diabetes is the most common cause of end-stage renal failure, this complication is considered the most dangerous and costly and is also one of the causes of death. Retinopathy is dangerous especially the majority of patients with diabetes do not notice any symptoms until eye suddenly lost vision. Then, despite aggressive treatment is costly and the ability to preserve your eyesight is very limited, the majority of patients will be blind permanently. Neuropathy is the cause of the removed extremities. The first histological lesions occur

early, but the clinical manifestations appeared only about 10 to 15 years so that early diagnosis of diabetic complications can cause tremendous value to patients and society. Computer-based methods are increasingly used to improve the quality of medical services. Artificial Intelligent focusing on creating systems that considered intelligent like humans. This type of system can minimized the dependence on the human expert. The proposed system for dealing with the problem of Diabetic Microvascular Complication Diagnosis is expert system. ESDMCD may be used by the limited clinical experience doctors to assist in decision-making in situations without consultation directly to a specialist.

In this work, methods and techniques related to the construction expert system will be presented. In the construction process an expert system, knowledge representation plays an important role. Therefore in this paper, we also will present a new model of knowledge representation suitable for representing knowledge in diabetic microvascular complication diagnosis domain.

The rest of this paper is organized as follows: Expert System for Diabetic Microvascular Complication Diagnosis (ESDMCD) is introduced in the section 2. The section 3 describes our design method. Modeling knowledge base is mentioned in the section 4, in this part we also proposes the KBCO-ADAPT model and presents the knowledge representation in diabetes microvascular complication diagnosis domain by using this model. The section 5 represents diagnostic problem and deductive method. The section 6 shows the testing results. And finally we give some conclusions and perspectives in the section 7.

2. Expert System for Diabetic Microvascular Complication Diagnosis (ESDMCD)

2.1 Expert System

In the book [9], the author represented that the expert system represents a branch of artificial intelligence, aiming to take the experience of human specialists and to transfer to a computer system. Specialty knowledge is stored in the computer. They are reasoned and the system derives specific conclusions by an execution system (inference engine). The expert system represents expertise knowledge as data or rules within the computer. It derives conclusions by running the knowledge base through an inference engine; a software program interacts with the user to process the results from the rules and data in the knowledge base. These rules and data can be called upon when needed to solve problems.

Diabetes is a very common serious disease in the modern world. It affects many organs in the body like heart, eyes, kidney, skin, nerves, blood vessels, foot etc. Complications of diabetes can cause serious consequences affecting the quality of life. Though these complications can not possible to cure completely, it can be well managed.

Early diabetic complication detection plays a crucial role because it helps diabetic patient improve the quality of life and reduce the serious consequences. Thus a diagnosis support expert system helping detect early the diabetic complications have practical significant.

2.2 ESDMCD

After collecting in-depth diabetic complication knowledge from endocrinologist at Go Vap Hospital, we represented this knowledge by using KBCO-ADAPT model to design the ESDMCD. This system uses deductive rules and forward chaining inference technique.

The purpose of the ESDMCD is helping and supporting user's reasoning but can not replace human experts. This system offers to the inexperienced user a diagnostic solution when human experts are not available.

The ESDMCD provides a user interface to interact with user, menu driven environment. In each diagnostic case, the system will ask a bunch of questions about the symptoms to user and user should give yes or no answer. According to these answers, the system will give to user the list of para clinical tests which system requires patient performing. After the para clinical test results are

recorded into the program, ESDMCD will carry out a diagnosis and provide the treatment method for each type of complication.

The screenshot shows a web-based diagnostic interface titled "KẾT QUẢ CHẨN ĐOÁN". It contains the following information:

- Họ và tên: Ngô Văn Châu
- Điện thoại: 071999 4 27 21 PM
- Ngày sinh: 07/1999 4 27 21 PM
- Email: Chau.Ngo@gmail.com
- Địa chỉ: 12 Đoàn Văn Bơ F 3. Quận 4
- Tiền Sử Đái Thào Đường:
- Loại: Loại 2

Lab results:

- Huyết áp tâm trương: 80
- Huyết áp tâm thu: 140
- HbA1C: 8
- Thời gian mắc bệnh: 1

Loại bệnh	Tên bệnh lâm sàng	Trả lời	Giá trị
Bệnh mắt	Nhìn mờ	<input checked="" type="checkbox"/>	
	Đau Mắt	<input type="checkbox"/>	
	Mù đỏ ngọt	<input type="checkbox"/>	
	Sợi dây mắt thấy xuất huyết võng mạc	<input checked="" type="checkbox"/>	
	Sợi dây mắt thấy tăng sinh võng mạc	<input checked="" type="checkbox"/>	
	Chụp dây mắt thấy xuất huyết võng mạc	<input checked="" type="checkbox"/>	
	Chụp dây mắt thấy tăng sinh võng mạc	<input checked="" type="checkbox"/>	

Kết quả: Bệnh võng mạc tăng sinh, biến chứng mắt đái tháo đường

Fig. 1 The diagnostic screen of ESDMCD.

This system has the structure of an expert system including six categories:

- The Knowledge Base contains the knowledge for solving some problems in diabetic complication diagnosis domain.
- The Inference engine will use the knowledge stored in knowledge base to solve problems. It must identify problem and use suitable deductive strategies to find out right rules and facts for solving the problem.
- The Working memory stores the facts and rules in the process of searching and deduction.
- The Explanation component supports to explain the phases, concepts in the process of solving the problem.
- The knowledge manager aims to support updating knowledge into knowledge base. It also supports to search the knowledge and test consistence of knowledge.
- The User Interface component of the system is required to have a specification language for communication between the system and users.

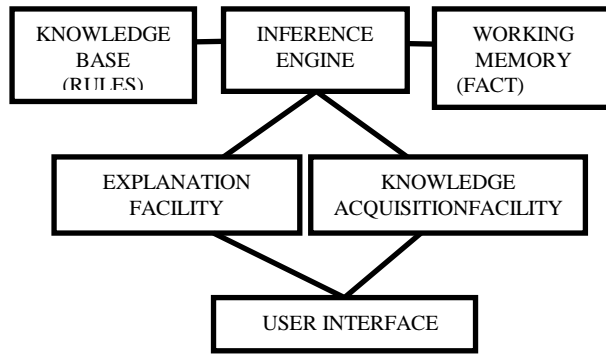


Fig. 2 Structure of the ESDMCD.

3. Design Method

In this section, we will present a process to construct ESDMCD based on the knowledge base and KBCO-ADAPT model. Besides, designed techniques in each phase will be presented.

This system is analyzed and designed with the following steps.

Step 1:

- Determine the domain knowledge.
- Determine the scope of the problem.
- Collecting real knowledge from expert in the diabetic domain.

Step 2:

- Designing the knowledge base.
- Representing knowledge in the step 1 and modeling knowledge based on the KBCO-ADAPT model.
- Designing algorithms.
- Establishing knowledge bases for the system based on KBCO-ADAPT model. Knowledge base can be organized to store by structured text files. They include the files below.
 - File ATT.xml stores name of concepts.
 - File RELATIONS.xml stores the specification of relations (The R component of KBCO-ADAPT model).
 - File CONCEPT.xml stores the concept of objects (The C component of KBCO-ADAPT model).
 - File RULES.xml stores deductive rules.
 - File PATIENT.xml stores certain objects.

Step 3

- Research on strategies for deduction to solve problems on computer. The most difficult thing is modeling for experience, sensible reaction and intuitional human to find the heuristics rules,

which were able to imitate the human thinking for solving problems

- Design the inference engine.
- Collecting issues, performance problems.
- Choosing inference engine and the design of deductive algorithms for solving problems.

Step 4:

- Designing the user interface of the system.
- Programming the software by using C# programming language and XML.

Step 5: Install.

Step 6: Develop a sample to test.

Step 7: In this stage we do testing, maintaining and developing the application.

4. Modeling Knowledge Base

Through research we recognized that the model for knowledge bases of computational objects (KBCO) in [5] are suitable cause this model is established from Object-Oriented approach to represent knowledge therefore it is useful to represent the medical knowledge. In medical domain, each disease has its own symptoms and para clinical tests that can be described fully and explicitly by KBCO model. In addition this model also use Rules component to represent the deductive rules similar the rules are used to diagnose by doctor.

However some components in this model should be changed to represent knowledge in medical domain more comprehensive and more explicit. In this paper, the new knowledge representation model based on KBCO model will be presented. This new model is also used for designing and constructing the ESDMCD.

The clinical process of diagnosing physicians showed that almost doctors rely on the clinical symptoms and the results of Para clinical tests to diagnose. In addition, doctors use the specific symptoms for differential diagnosis. Beside the classic knowledge in the medical literature, the doctors also use a lot of practical experience to definitive diagnosis. Therefore the knowledge in this domain is really complex.

We studied many knowledge representation models to represent this kind of knowledge. Finally, we recognized that the model for knowledge bases of computational objects (KBCO) mentioned in [5] with some changes in components is appropriate to represent the knowledge in diabetic complication diagnosis domain. Components such as C, R and Rules of KBCO model can be used to describe knowledge in the medical diagnosis because in this area there are many objects including many attributes

such as some complications of diabetes. They have also relationships together. Besides, many deductive rules are used to diagnose by doctors.

However, H, Ops components are not necessary because the knowledge in the medical diagnosis domain is not required to perform calculations and there is not a special relationship with H component.

Finally, if only using the C, R and Rules components can not describe the knowledge completely so that Attr and Patient components are added. Attr component is a set of base attributes. It is used to describe all symptoms of diabetic complications. Patient component includes a set of attributes and a set of concepts of objects. This component is used to describe the diagnosed patient. The adapted model based on KBCO is called KBCO-ADAPT.

4.1 Model for Knowledge Bases of Computation Objects (KBCO)

The model for knowledge bases of computational objects (KBCO) mentioned in [5] consists of six components:

(C, H, R, Ops, Funcs, Rules)

These components are defined as follows

- **C** component is a set of concepts of computational objects. Each concept in C is a class of Com-objects.
- **H** component is a set of hierarchy relation on the concepts.
- **R** component is a set of relations on the concepts.
- **Ops** component is a set of operators.
- **Funcs** component is a set of functions.
- **Rules** component is a set of rules.

There are relations represent specializations between concepts in the set C; H represents these special relations on C. This relation is an ordered relation on the set C, and H can be considered as the Hasse diagram for that relation.

The set R is a set of other relations on C, and in this case a relation r is a binary relation it may have properties such as reflexivity, symmetry, etc....

The set Ops consists of operators on C. This component represents a part of knowledge about operations on the objects.

The set Funcs consists of functions on Com-Objects. Knowledge about functions is also a popular kind of knowledge in almost knowledge domains in practice, especially fields of natural sciences such as fields of mathematics, fields of physics.

The set Rules represents for deductive rules. The set of rules is a certain part of knowledge bases. The rules represent for statements, theorems, principles, formulas, and so forth. Almost rules can be written as the form “if <facts> then <facts>”. In the structure of a deductive rule, <facts> is a set of facts with certain classification. Therefore, deductive rule is used in the KBCO model. Facts must be classified so that the component Rules can be specified and processed in the inference engine of knowledge base system or intelligent systems.

4.2 The KBCO-ADAPT Model

KBCO-ADAPT model shows that the medical diagnostic knowledge can be represented completely.

The KBCO-ADAPT model consists of five components:

(Attr, C, R, Rules, Patient)

The meaning of the components is as follows:

- **Attr** component is a set of base attributes.
- **C** component is a set of concepts of objects. Each concept in C is a class of objects.
- **R** component is a set of relations on the concepts.
- **Rules** component is a set of deductive rules on facts.
- **Patient** component is a set of attributes or objects.

In this part, the knowledge about Diabetic Microvascular Complication are represented by using KBCO-ADAPT Model

• **Attr component:**

Attr component is the list or a set of base attributes: Attr = { x_1, x_2, x_n } in which every attribute has a value in the fixed value region as float, integer, Boolean, string.

This component is used to describe all symptoms of diabetic complications and the symptoms used to distinctive diagnosis.

Example:

```
Attr= {  
BloodGlucoseTest      : float;  
DiabetesType           : string;  
HbA1c                  : float;  
BlurryVision          : Boolean;  
Pain                   : Boolean;  
Numbness               : Boolean;  
Soreness               : Boolean;  
HighbloodpressureHis  : Boolean;  
Numb                   : boolean;  
Retinopathy            : string;  
Nephropathy           : string;  
Neuropathy            : string;  
DiastolicBlood Pressure : integer;
```

SystolicBlood Pressure : integer
 ... }

• **C component is a set of concepts of objects**

C includes some concepts of objects as follows:

Basic concept: A basic concept in C is a class of objects that has structure including a set of attributes $(O) = \{x_1, x_2, x_k\} \in Attr$. This class of objects can be used to describe the “level 1” objects.

The basic concepts can be used to describe a symptoms list or a para clinical test list of a diabetic complication.

There are some basic concepts in ESDMCD system as follows.

- Retinopathy_PhysicalSymptomsList,
- Retinopathy_FunctionalSymptomsList,
- Retinopathy_ClinicalTestList,
- Nephropathy_FunctionalSymptomsList,
- Nephropathy_PhysicalSymptomsList,
- Nephropathy_ClinicalTestList,
- Neuropathy_FunctionalSymptomsList,
- Neuropathy_PhysicalSymptomsList,
- Neuropathy_ClinicalTestList

Example:

```
Attr(Retinopathy_PhysicalSymptomsList)
{Glaucoma : Boolean;
Cataracts : Boolean;
Ophthalmoscopy_RetinalHemorrhage : Boolean;
Ophthalmoscopy_RetinalVascularProliferative: Boolean;
Ophthalmoscopy_Macularedema : Boolean }
```

The class of “level 1” objects

The class of “level 1” objects is a class of objects having structure include a set of attribute $Attr(O) = \{x_1, x_2, \dots, x_k\} \in Attr$ and $A_1(O)$ has structure including a set of attributes whose data type is the basis concept. These objects can be used to define the “level 2” objects.

In ESDMCD, the class of “level 1” objects is used to describe a type of diabetic complication such as NeuropathyComplication_Diabete, NephropathyComplication_Diabete, RetinopathyComplication_Diabete.

Example:

The RetinopathyComplication_Diabete object has structure including a set of attributes and basic concepts.

```
Attr (RetinopathyComplication_Diabete) =
{EyesInjury : Boolean;
Keratitis : Boolean;
BlindnessHistory : Boolean;
Retino_FunctionalSymptoms :
Retinopathy_FunctionalSymptomsList;
Retino_PhysicalSymptoms :
Retinopathy_PhysicalSymptomsList;
```

```
Retino_ClinicalTest :
Retinopathy_ClinicalTestList }
```

The class of “level 2” objects is a class of objects has structure including a set of attributes $Attr(O) = \{x_1, x_2, \dots, x_k\} \in Attr$ and $A_1(O)$ has data type is the class of “level 1” objects. In ESDMCD, the class of “level 2” objects is used to describe microvascular complication of diabetes.

Example:

The MicrovascularComplication_Diabete object has a set of “level 1” objects as follows:

```
Attr(MicrovascularComplication_Diabete)=
{ HighbloodpressureHis : Boolean;
DiabetesType : string;
Diabetic_Retinopathy :
RetinopathyComplication_Diabete;
Diabetic_Nephropathy :
NephropathyComplication_Diabete;
Diabetic_Neuropathy:
NeuropathyComplication_Diabete }
```

• **R is a set of relations on C**

Each relationship is determined by

[<name of relation>, < object1 >, < object2 >]

Relation “belong to” of an object and an object

Example :

```
Obj1: Retinopathy_PhysicalSymptomsList
Obj2: RetinopathyComplication_Diabete
[“belong to”, Obj1, Obj2]
```

• **Rules**

Rules relating to the objects are a set of deductive rules on the facts. Each rule described in this rules represents the inference rules in order to find out the new fact from the previous known facts. In general, each rule consisting of two parts: the hypothesis and conclusion. Almost rules can be written as the form “if <facts> then <facts>”. In the structure of a deductive rule, <facts> is a set of facts with certain classification. Facts must be classified so that the Rules component can be specified and processed in the inference engine of intelligent systems. Each rule is described as $\{fact1, fact2, fact3, factn\} \Rightarrow \{ fact1, fact2, fact3, \dots factm\}$.

There are many rules in the system. We represented these rules in the Rules.xml file. This is the structure of a rule used to diagnose in the ESDCMD.

```
<begin-rule>
<hypothesis_part>
</hypothesis_part>
<goal_part>
</goal_part>
<compare_value>
</compare_value>
</begin-rule>
```

Example:

```
<begin-rule>
  <hypothesis_part>
    DiabetesType                : type1;
    BlurryVision                : true;
    Diabetes Duration            : -&gt;-5 ;
    Ophthalmoscopy_RetinalHemorrhage : true;
    Ophthalmoscopy_RetinalVascularProliferative: true;
    Ophthalmoscopy_Macularedema   : true;
  </hypothesis_part>
  <goal_part>
    Retinopathy: Diabetic Retinopathy Complication
  </goal_part>
  <compare_value>
    Diabetes Duration -&gt;-5
  </compare_value>
</begin-rule>
```

To apply the rules described in the knowledge base we have the following fact types.

Fact about an object type

This kind of fact informs about the object type.

We show this fact type by structured list:

[<object>, <object type>]

Example:

[Obj, "RetinopathyComplication_Diabete"]

In this system we have some objects such as:

[obj1, MicrovascularComplication_Diabete],
 [obj2, RetinopathyComplication_Diabete],
 [obj3, Retinopathy_PhysicalSymptomsList],
 [obj4, Retinopathy_FunctionalSymptomsList],
 [obj5, Retinopathy_ClinicalTestList],
 [obj6, NephropathyComplication_Diabete],
 [obj7, Nephropathy_FunctionalSymptomsList],
 [obj8, Nephropathy_PhysicalSymptomsList],
 [obj9, Nephropathy_ClinicalTestList],
 [obj10, Neuropathy Complication_Diabete],
 [obj11, Neuropathy_FunctionalSymptomsList]

The fact for the determination of an object

This fact is used to identify an object (attributes considered as known) or an attribute of an object by a value.

[<object> | <object>.< attribute value >]

Example:

Obj: Retinopathy_PhysicalSymptomsList

Obj.Ophthalmoscopy_RetinalVascularProliferative :true.

[<Obj> | <Obj>.< attribute value >]

The "Ophthalmoscopy_RetinalVascularProliferative" attribute belongs to Retinopathy_PhysicalSymptomsList object. The attribute has value "true".

The fact about a relationship

This fact type performs a relationship between the objects or attributes of objects. This fact type can be represented by the structured t:

[<relationship fact >, < Obj1>, < Obj2>]

Example:

Obj1: RetinopathyComplication_Diabete.

Obj2: Retinopathy_PhysicalSymptomsList.

[<belongto>, < Obj1>, < Obj2>]

• Patient component

Patient component is a class of objects has structure including a set of its own attributes, a set of attributes $Attr(O) = \{x_1, x_2, \dots, x_k\} \in Attr$ and a set of objects $A1(O) \in C$. This component is used to describe the diagnosed patient.

Patient =

```
{PatientName                : string;
  YOB                       : integer;
  Address                   : string;
  PrehistoricDiabetes       : Boolean;
  Diabetes Duration         : float;
  BloodGlucoseTest         : float;
  DiabetesType              : string;
  HbA1c                     : float;
  BlurryVision              : Boolean;
  Numbness                  : Boolean;
  Soreness                  : Boolean;
  Retinopathy               : string;
  Nephropathy               : string;
  Neuropathy                : string;
  DiastolicBlood Pressure   : integer;
  SystolicBlood Pressure    : integer;
  Diabetic_Retinopathy     :
    RetinopathyComplication_Diabete;
  Diabetic_Nephropathy     :
    NephropathyComplication_Diabete;
  Diabetic_Neuropathy      :
    NeuropathyComplication_Diabete;
[obj1, MicrovascularComplication_Diabete],
[obj2, RetinopathyComplication_Diabete],
[obj3, Retinopathy_PhysicalSymptomsList],
[obj4, Retinopathy_FunctionalSymptomsList],
[obj5, Retinopathy_ClinicalTestList],
[obj6, NephropathyComplication_Diabete],
[obj7, Nephropathy_FunctionalSymptomsList],
[obj8, Nephropathy_PhysicalSymptomsList],
[obj9, Nephropathy_ClinicalTestList],
[obj10, Neuropathy Complication_Diabete],
[obj11, Neuropathy_FunctionalSymptomsList],
[obj12, Neuropathy_PhysicalSymptomsList],
[obj13, Neuropathy_ClinicalTestList]...}
```

5. Diagnostic Problem and Deductive Method

5.1 Diagnostic Problem

General components, which are present in a diagnostic procedure in most of the various available methods include:

Complementing the already given information with further data gathering, which may include questions of the medical history (potentially from other people close to the patient as well), physical examination and various diagnostic tests.

A diagnostic test is any kind of medical test performed to aid in the diagnosis or detection of disease. Diagnostic tests can also be used to provide prognostic information on people with established disease.

Processing of the answers, findings or other results. Consultations with other providers and specialists in the field may be sought.

Based on the knowledge base of the microvascular complication of Diabetes is modeled COKB-ADAPT model above.

Input = The set of symptoms are provided by patient, symptoms are detected by the doctor on physical examination and the results of diagnostic tests.

Output = Diagnostic conclusion with
 Diabetic Neuropathy Complication,
 Diabetic Retinopathy Complication,
 Diabetic NephropathyComplication or not?

Example :

Input:

The set of symptoms are provided by patient

DiabetesType: type1;

Diabetes Duration: 7;

The set of symptoms are detected by the doctor on physical examination

Ophthalmoscopy_RetinalVascularProliferative: true;

Results of laboratory test

MicroalbuminNieu: 45

Output: Diagnosis

Nephropathy : Diabetic Nephropathy Complication;

5.2 Clinical Diagnostic Process

In beginning, patient reports the primary symptoms to a physician, and then he (or she) can detect more symptoms on physical examination. If the probability of certain diagnoses is halfway between likely or unlikely, the physician can request certain tests on the patient such as blood test, Micro albumin...ect.

Like the history and physical exam, diagnostic testing provides new information that further narrows the differential diagnosis.

Finally based on the clinical symptoms and the results of para clinical test, the physician summarizes the most likely diagnosis.

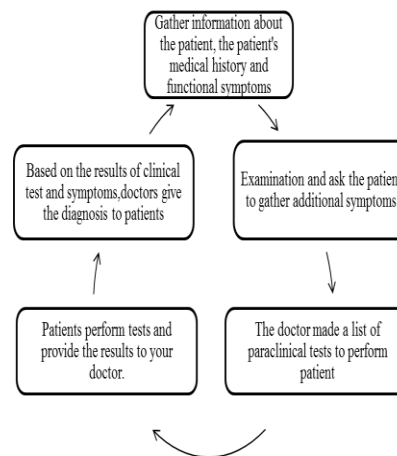


Fig. 3 Actual diagnostic process

5.3 Deductive Method

Inference engine used in the expert system for diabetic complication diagnostic is forward chaining method. This deductive method is often used in expert system. The solution process is started with the collection of information and this information is deduced to draw conclusions.

The inference strategy begins with collective known facts, new facts learned by using the law assumes that the fit known facts, and continue this process until you see the status of the target, or for until no rule that matched the event is known or inferred events.

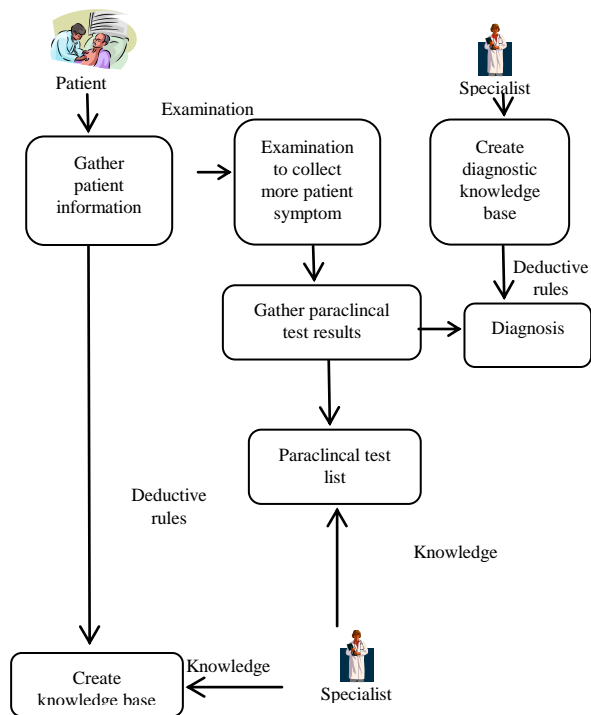


Fig. 4 Diagnostic process of ESDMCD

6. The Testing Result

Data collected from 4 District General Hospital from July 16, 2012 to November 12, 2012 is used to evaluate the expert system for diabetes microvascular complication diagnosis (ESDMCD). This is the actual data for research on the diabetic microvascular complications of a doctor working at District 4 General Hospital.

The ESDMCD be assessed by comparing the diagnostic results of expert physicians with the expert system in 106 patients.

Comparison of the diagnostic expert system ESCMD and doctor shows cases of 106 cases with 2 different diagnostic results. 104 cases resulted in equivalent proportion of 98.11%.

A proper diagnose case of system

PatientName: Nguyen thi Nguyet Anh
 YOB :1963
 Address: 12 Doan Van Bo Street, 4th District
 PrehistoricDiabetes : true;
 Diabetes Duration : 6;
 HbA1c : 12,5;
 BlurryVision : True;

```

    BloodGlucoseTest      : 12.5;
    DiabetesType          : type1;
    BlurryVision          : true;
    Soreness               : true;
    Diabetic_Retinopathy  :
        RetinopathyComplication_Diabete;
    Diabetic_Nephropathy  :
        NephropathyComplication_Diabete;
    Diabetic_Neuropathy   :
        NeuropathyComplication_Diabete;
    [obj1, MicrovascularComplication_Diabete],
    [obj2, RetinopathyComplication_Diabete],
    [obj3, Retinopathy_PhysicalSymptomsList],
    [obj4, Retinopathy_FunctionalSymptomsList],
    [obj5, Retinopathy_ClinicalTestList],
    Obj1.obj4.Ophthalmoscopy_RetinalHemorrhage =true;
    Obj1.obj4.Ophthalmoscopy_RetinalVascularProliferative
    : =true;
    Obj1.obj4.Ophthalmoscopy_Macularedema     = true;
    
```

Conclusion:

Diagnosis of ESDMCD: Diabetic Retinopathy
 Complication.
 Diagnosis of physican: Diabetic Retinopathy
 Complication

7. Conclusions and Perspective

The expert system for diabetes microvascular complication diagnosis (ESDMCD) is an application can help the doctors with limited experience in providing decision support system, interactive training tool and expert knowledge. This work proposed the KBCO-ADAPT model to represent the knowledge in this domain.

The ESDMCD imitated diagnostic process of doctor in clinical practice. By using exact rules, the system gave the relatively accurate diagnostic results similar doctor's diagnosis with proportion 98.11%. After testing the system, the specialist gave us the positive feedback. Therefore this system is appreciated very high in the practical applications by endocrine specialists. In the future, we will continue to research the COKB-ADAPT model to complete the program for solving problems.

The KBCO-ADAPT also proposed a natural way for the knowledge representation in Diabetic Microvascular Complication Diagnosis domain. This model is also a tool to design practical knowledge bases and algorithm to solve medical diagnosis problems.

Acknowledgements

We gratefully acknowledge Endocrine Specialist Minh Chau Tran Hoang for providing her expertise knowledge and Endocrine Specialist Chuong Phuong Tran Thi for using her data to evaluate system in this work.

References

- [1] George F. Luger: William A Stubblefield Artificial Intelligence. Addison Wesley Longman, Inc. (1998).
- [2] Leszek Rutkowski: Computational Intelligence: Methods and Techniques. Springer-Verlag Berlin Heidelberg (2008).
- [3] Michel Chein, Marie-Laure Mugnier: Graph-based Knowledge representation: Computational foundations of Conceptual Graphs. Springer-Verlag London Limited (2009).
- [4] Trach The Mai, Khue Thy Nguyen: Biện chứng mãn tính của bệnh Đái Tháo Đường. Nội tiết học đại cương, Y Học Publisher (2007).
- [5] Nhon Do: Model for Knowledge Bases of Computational Objects, International Journal of Computer Science Issues, vol 7, issue 3, No 8, pp 11-20 (2010).
- [6] Nhon Do, Hien Nguyen: Model for Knowledge Representation using Sample Problems and Designing a Program for automatically solving algebraic problems. World Academy of Science, Engineering and Technology, (ICEEEL), Paris (2010).
- [7] Nhon Do, Hien Nguyen: A reasoning method on Computation Network and Its applications. 2011 International MultiConference of Engineers and Computer Scientists, IMECS 2011, ISBN: 978-988-18210-3-4, pp.137-141, Hongkong, March (2011).
- [8] Trach The Mai, Khue Thy Nguyen: Nội tiết học đại cương. Y Hoc Publisher (2007).
- [9] Petrică Vizureanu: Expert Systems. Intech (2010).
- [10] Ronald J. Brachman, Hector J. Levesque: Knowledge Representation and Reasoning. Elsevier, Inc. (2004).
- [11] Randall Davis, Howard Shrobe, Peter Szolovits: What is a Knowledge Representation? AI Magazine Volume 14 Number 1 AAAI, SPRING (1993)
- [12] Samy S. Abu Naser, Abu Zaiter A. Ola: An Expert System for Diagnosing Eye Diseases Using Clips. Journal of Theoretical and Applied Information Technology, P 923, JATIT (2008).



Truc Ly Thi NGUYEN is currently a lecturer in the Department of Computer Science in Industrial University of Ho Chi Minh, Vietnam. She got her BA in medical in 1992 from Medical and Pharmaceutical University of Ho Chi Minh City, Viet Nam. She graduated her bachelor degree in Computer Science from University Technology of Ho Chi Minh City, Vietnam, in 2005. Her research interests include Artificial Intelligence, computer science, and their practical applications, especially intelligent systems, expert system, knowledge base, and knowledge systems.



Nhon Van DO is currently an associate professor in the faculty of Computer Science at the University of Information Technology, Ho Chi Minh City, Vietnam. He got his MSc and Ph.D. in 1996 and 2002 respectively, from The University of Natural Sciences –National University of Ho Chi Minh City. His research interests include Artificial Intelligence, computer science, and their practical applications, especially intelligent systems and knowledge base, knowledge systems.

Implementation of Information and Communication Technology in Inclusive Education in Nigeria: Future Perspective

1. Eskay, M., Ph.D.
University of Nigeria, Nsukka

2. Ezegbe, N., Ph.D.
University of Nigeria, Nsukka

3. Anyanwu, J., Ph.D.
University of Nigeria, Nsukka

4. Ikwumelu, S. N., Ph.D.
Ebonyi State University

Abstract

Integration of learners with and without disabilities into Nigerian educational system had been on the track few decades back. Satisfaction or dissatisfaction on implementation of inclusive education can only be achieved through today's understanding of Information and Communication Technology (ICT). Future trends on implementation of inclusive education will have more meaning when various strategies are sought and are given priority. Awareness is high everywhere in the world that human success and prosperity can only be achieved through new ideas, new designs and innovations which are associated or anchored through collaborative, cooperative and consultative partnership. It has therefore become imperative for government to begin to embrace a new paradigm shift in the training of teachers and education of learners with disabilities in this technological age.

Key words: ICT, *litigation, collaboration, future perspective, implementation, inclusive education and creativity.*

1. Introduction

Education generally is a process of preparing individuals for sound and effective social living. Education embodies the major Nigerian values of freedom; equity and cohesion. Education needs to reach out to the entire human capital of a country and must not reject any talent, no matter how hard it is to encourage and no matter how peripheral it may seem. This is due to the fact that unequal access to education technology is an incalculable loss of human potential and threatens social cohesion. It is a means of social mobility.

Generally in Nigeria, Education through the use of Information and Communication Technology (ICT) has been mainly for the normal children with few schools for the challenged or disabled children. Currently, there has been the realization that the disabled children need to be equally educated through the use of ICT because the few schools in existence do not cater adequately and effectively for about half a million disabled persons in Nigeria out of which one hundred and seventy thousand, five hundred and seventy seven (170,577) are of school age (Alao 2006). Thus, Education must be planned and organized to cater for all categories of children irrespective of their abilities and disabilities (Taura, 2011). This planning entails the application of ICT training in inclusive education setting.

2. Concepts Surrounding Inclusive Education

Inclusive education is of paramount significance in achieving the goals of Education for All (EFA). It involves pursuing excellence without sacrificing integration, offering all students equal opportunities in educational computer technology regardless of their difficulties, expecting the best of each student and developing their full potential, while taking special needs into account. Expanding further, Acedo, Operti, Brady and Duncombe (2011), stated that inclusive education involves the conversion of schools into learning communities that could foster a sense of inclusion and mutual support. To achieve this, schools have to recognize that every person, students, teachers, parents, school staff and members of civil society with whom they interact, is a valued member of the community and have not only roles to play but also talents that can help in ensuring the success of all its components irrespective of whether they are physically disabled or have high skills. Contributing, Okuoyinbo (2001) asserted that ICT in inclusive education means bringing students with disabilities regardless of the nature and severity of their disabilities into all activities of general education such as schools, classroom and the surrounding communities. In effect, ICT in inclusive schools integrates, enrich and celebrate diversity and intercultural difference.

Ademokoya and oyewumi (2001) affirmed that an inclusive school is an educational institution in which all available resources, including the use of ICT are collaboratively utilized to meet the educational needs and challenges of all children who reside in its catchment area. According to this, Acedo and Operti et al (2011) stated that inclusive education has to be developed in close cooperation with all authorities, institutions, associations and civil society, as

well as embedded into all policy areas. They further stated that participation is not only essential and necessary, but also a matter of rights and essential to national development.

Afoi, Saad and Atukum (2008) had earlier expressed the opinion that the use of ICT in inclusive education is essentially a program that enables all learners to participate fully in the life and work of mainstreamed settings to meet the learners' needs through technological innovation. They went further to say that inclusive education is a continuous process of breaking down barriers to learning and active participation for learners with disabilities. It can also be said that inclusive education is concerned with a system of education which gives equal opportunities for all learners to jointly undertake learning activities without discrimination. From these definitions it is clear that emphasis is on giving everybody in the society an equal opportunity to benefit from formal education. Placing the disabled children into the numerous public schools in Nigeria entails providing the needed technological equipment, support services and popularization of special education.

Special education is an aspect of education that appreciates and recognizes the needs of disabled people and makes allowances for the use of special equipment such as computer technology and methods of teaching according to individual needs. However, the educational system in Nigeria which gives room for segregation which is a system of separating persons with disabilities from their normal peers for the purpose of educating them and integration, has not only deprived many disabled persons the opportunity to be educated formally, but also hindered many normal persons from gaining relevant information about the potentials and capacities of persons with disabilities.

In order to legalize inclusion, Decree 3 of 1993 provides the rights and privileges for people with disabilities including Education, rehabilitation, health, vocational training, counseling and employment among others. The national policy on education (FRN, 2004) equally stipulated that the Federal ministry of education shall provide special education in collaboration with appropriate bodies for disabled people irrespective of type and severity. The aim of the special education include to integrate special classes into public schools, monitor people with disability so as to plan for their welfare, provide needed materials and ICT equipment that will enhance their learning.

3. Implementation of Inclusive Education

Special education started to gain recognition in Nigeria in the 70s during the regime of General Gowon as the military head of state (Ademuyiwa and Oseni, 2006). Following this, the federal government made provisions for special education in the National policy on education with the following sum-up objectives:

- To give meaning to the idea of equalizing educational opportunities
- To provide education for all disabled children through the use of ICT
- To provide opportunities for exceptionally gifted children to develop at their own rate.

This policy suffered a setback because of government's inability to fulfill recommendations in the policy. For instance, the suggested ten percent (10%) of funds for programs in education required to be spent on persons with disability was not provided for the purpose. Besides, the ten percent (10%) of Employment in the public sector expected to be reserved for the persons with disability is on paper only. The establishment of the national commission for people with disability is yet to be implemented. The problem of finance and lack of special education commission for persons with disabilities has hindered inclusive education program in Nigeria. Efforts were made by government to expand facilities for the education and training of disabled persons especially in primary schools while in the secondary school, integration is mostly vogue. Part of the achievement is the promulgation of Degree 3, of 1993 which gave legal backing to the demands and mandates of previous special education policy.

4. Future Perspectives: Shifting Paradigm

Through the years, implementation of ICT in inclusive education in Nigeria has come a long way; however, it still has a long way to go. For example, the only special education mandate comes from Section 8 of the National Policy on Education. As a result, accountability for inclusive education is affected and services for people with disabilities are not reachable. For instance there is no mobility cane for the blind, talking watch, typewriter and audiovisual equipment to mention a few. Counseling services are not provided to help early identification of disabled children, and enhance the adjustment of those already identified. Consequently, these learners are left without much assistance in learning. Disability can be identified through observation of the children both inside and outside the classroom setting, scrutinizing school results, consulting the teachers and

individual counseling with the child. Sometimes, tests are administered to complement the non-test results obtained. The counselor will then be able to facilitate immediate intervention strategies that can influence the adjustment of the identified disabled person. JOHN D. KRUMBOLTZ, THIERRY G. KOLPIN School counselors help to make learning a positive experience for every student. They are sensitive to individual differences. They know that a classroom environment that is good for one child is not necessarily good for another. Counselors facilitate communication among teachers, parents, administrators, and students to adapt the school's environment in the best interests of each individual student. They help individual students make the most of their school experiences and prepare them for the future. The duties and roles of many counselors began to change considerably. Counselors started finding themselves as gatekeepers to Individualized Education Programs (IEP) and Student Study Teams (SST) as well as consultants to special education teachers, especially after passage of the Americans with Disabilities Act in 1990. School counselor now has a working relationship with students' families and with community social agencies. Although the roles of school counselors vary among settings, common tasks include individual counseling, small-group counseling, large-group or classroom presentations, involvement in school wide behavior plans for promoting positive and extinguishing negative behaviors, and consulting with teachers, parents, and the community. Additional duties might include developing classroom management plans or behavior plans for individual students, such as conducting SST and IEP meetings. Without a clear understanding of the various socio-economic and cultural characteristics of diverse learners, it is difficult to evolve strategies and develop plans at classroom, school and system levels to teach children from diverse backgrounds. It is therefore necessary that relevant data and information on diverse learners is collected, examined and analyzed in order to inform and shape policies and practices to make 11 classrooms and schools inclusive and responsive to the learning needs of children from diverse backgrounds.

Child profiles can be an important and effective tool to promote inclusive classrooms. There is first a need to recognize the changing social composition of learners in the classroom resulting from an inflow of children from diverse backgrounds in terms of caste, class, gender, ethnicity, language and religion. This increased diversity presents new issues and challenges in curriculum design, teaching-learning practices and processes, learning materials, and teacher education that

meet the different learning needs of these children. These issues and challenges cannot be addressed unless they are first recognized by policymakers and practitioners. Read more: <http://www.answers.com/topic/school-guidance-and-counseling#ixzz2WDVNQHGJ>, read more: <http://www.answers.com/topic/school-guidance-and-counseling#ixzz2WDU7wWcy>. To look at the future, Nigeria must shift its paradigms and join other progressive nations to advocate for the rights of learners with disabilities and help them join the mainstream society (Mukuria & Obiakor, 2004).

From the researchers' perspective, Nigeria must build on the foundational knowledge that it has established. Proactive efforts must be made to:

- *Enforce a national policy on inclusive education*—currently; there are no laws that enforce a national policy on special education. Unlike other countries that have special education laws that guide their special education activities, Nigeria still struggles with this issue. Even the Section 8 of the National Policy does not provide room for due process. The lack of legal enforcement on special education makes it difficult for people with disabilities and their parents to fight for their rights and gain public support.
- *Institute legitimate advocacy groups*—an institution of advocacy groups will be extremely helpful. Individuals must lobby for the rights of persons with disabilities. Additionally, parent advocacy must be established.
- Telephone communication devices for the deaf (TDDs) should be procured for students use.
- Purchase of other hearing enhancement systems for the use of the deaf patrons who do not have.
- Maintain an extensive collection of books and information for people with disabilities
- *Recognize and accommodate various forms of exceptionalities* - At present, some exceptionality such as blindness and deafness are recognized in Nigeria. There is need to broaden the scope.

- *Give quality and equal educational opportunities to all learners* -All learners deserve quality and equitable treatment and education. The local, state, and federal governments must be involved in these processes.
- *Encourage positive societal attitudes to disability and the disabled* – In all human societies, people value respect and positive attitudes. If the government of Nigeria could institute public education on special education through media and other channels, it would go a long way in bringing respect to the special education community.
- *Be independent, think independently, and talk independently:* Nigeria’s independence must be reflected in her words and deeds. Not only does she need policies that depict her independence from colonial domination, but also Nigeria’s educational language must also be independent.

5. Conclusion & Recommendation

Implementation of Information and Communication Technology in inclusive education in Nigeria, like in many African countries that were former colonies, is a post-colonial phenomenon. In relative terms, it has remained rudimentary in its operation and limited in its functionality. Historically, inclusive education was instituted as an excellent path for providing equal education to all Nigerian citizens, a fact that has been more theoretical than pragmatic. Thus, it has found itself struggling with how best to make it better for learners who are disenfranchised. It is not surprising that for decades such struggle has not produced any measurable progress in the education and service delivery for all learners. By all standards, inclusive education has continued to flounder in mediocrity. To join other progressive countries in recognizing, protecting, and maintaining the rights of persons with and without exceptionalities, it is imperative that Nigeria begins to shift its paradigm in the twenty-first century by embracing and educating all learners through Information Technology and putting away the old tradition of negative perceptions on these learners. It is important that to institute ways to include all learners in Nigeria’s mainstream activity and help them maximize their fullest potential, as no nation has ever gravitated towards reaching its maximum potential without putting into consideration all the human capital with which it is endowed.

6. Future of Inclusive Education in Nigeria

In addition to the scholastic views, National policy on Education and decree 3 of 1993 supported inclusive education in Nigeria. Blue prints on special education also amplified some of the issues raised in the National policy on Education and made new proposals to the National Board for the Education of disabled persons through new technological innovations and gave legal basis to special education policies through appropriate and effective legislation (Ayoku 2007). This is to postulate that future inclusive education in Nigeria can be improved. It is obvious that things have not been working well in the field of special education. However, if mistakes made can be corrected along with the under listed recommendations, the hope of inclusive education is brighter.

- There should be serious effort to ensure that issues contained in the decree 3 of 1993 are implemented by relevant agencies without delay.
- Participation of relevant stakeholder in the execution of policies related to special education should be by merit.
- Telephone communication devices for the deaf (TDDs) should be procured for students use.
- Purchase of other hearing enhancement systems for the use of the deaf patrons who do not have.
- Maintain an extensive collection of books and information about deafness
- Curriculum content that will serve the needs of learners with and without disabilities should be designed by experts.
- there should be supervision by government agencies to reduce poor standards
- Counseling services recommended in the 1993 decree has to be established to organize and mount periodic programs both to raise the consciousness of people about persons with disabilities and help the disabled persons adjust to the normal school environment. It will also enhance the early identification of persons with disability so as to ensure early intervention and subsequent early completion of education at all levels as stipulated by the national policy on education.

Future of inclusive education will be better and encouraging when classrooms, good environment, enough funds is spent on special education and employment after each level of education is guaranteed. Besides, when the national commission for people with disabilities is established and

headed by disabled person who is believed will be capable of handling matters that related to disabilities, things will work better.

References

- Ademokoya, O.A. & Oyewumi, A.M. (2001). Applying inclusive education to the hearing impaired child in the nursery school. Ibadan, Journal of Special Education, University of Ibadan, 9 (1) 20-28
- Ademuyiwa, O.A. & Oseni, I.A. (2006). Special education in the 20th century: The journey so far in Nigeria. Ibadan Glory land Publishing Company
- Afoi, P.A., Saad, M.T., & Atukum, E. (2008). Inclusive education and its implications for teachers in Nigeria. The Journal of the National Council Exceptional Children, 10 (2b), 18-36
- Alao, K. (1999). Guidance and counseling: Implication for parents/guidance of children with special needs. A paper presented at the International Conference on Special Children Empowerment, Ibadan, Oyo State
- Ayoku, F.A. (2007). Implementation of special education policy in Nigeria. In Collaborative Partnership in Special Education, Ibadan. Glory land Publishing Company.
- Federal Republic Nigeria (2004). National Policy on Education, NERDC Press
- Mukuria, G., & Obiakor, F.E. (2004). Special education issues and African diaspora. *Journal of International Special Needs Education*, 7, 12-17
- Okuyinbo, J.M. (2001). Inclusive : An alternative education placement for exceptional children in Nigeria's New Political Dispensation. In J.A. Ademokoya (Ed). Exceptional Nigerians in the New Political Dispensation. Ibadan, Option books.
- Taura, U.A. (2011). Inclusive education: Teachers preparation and its implication in the special education. Journal Publication of the Nigerian Association of Special Education Teachers, 10 (1), 1-10.
- Acedo, C, Operti, R, Brady, J and Duncombe, L (2011). Interregional and Regional Perspectives on Inclusive Education: Follow-up of the 48th session of the International Conference on Education. UNESCO: International Bureau of Education.
- UNICEF (2010) Report On Inclusive Classroom, Social Inclusion/Exclusion and Diversity: Perspectives, Policies and Practice.
- Deshkal Publication *KRUMBOLTZ, JOHN D & KOLPIN THIERRY G (eds) (2011) Gale Encyclopedia of Education School Guidance and counseling. Retrieved from <http://www.answers.com/topic/school-guidance-and-counseling#ixzz2WDWovveO>*

IJCSI CALL FOR PAPERS JANUARY 2014 ISSUE

Volume 11, Issue 1

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas. See authors guide for manuscript preparation and submission guidelines.

Accepted papers will be published online and indexed by Google Scholar, Cornell's University Library, DBLP, ScientificCommons, CiteSeerX, Bielefeld Academic Search Engine (BASE), SCIRUS, EBSCO, ProQuest and more.

Deadline: 10th December 2013

Online Publication: 31st January 2014

- Evolutionary computation
- Industrial systems
- Evolutionary computation
- Autonomic and autonomous systems
- Bio-technologies
- Knowledge data systems
- Mobile and distance education
- Intelligent techniques, logics, and systems
- Knowledge processing
- Information technologies
- Internet and web technologies
- Digital information processing
- Cognitive science and knowledge agent-based systems
- Mobility and multimedia systems
- Systems performance
- Networking and telecommunications
- Software development and deployment
- Knowledge virtualization
- Systems and networks on the chip
- Context-aware systems
- Networking technologies
- Security in network, systems, and applications
- Knowledge for global defense
- Information Systems [IS]
- IPv6 Today - Technology and deployment
- Modeling
- Optimization
- Complexity
- Natural Language Processing
- Speech Synthesis
- Data Mining

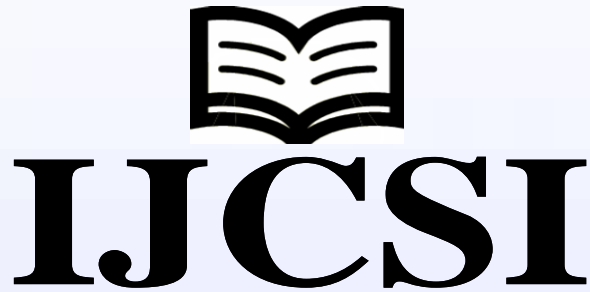
For more topics, please see <http://www.ijcsi.org/call-for-papers.php>

All submitted papers will be judged based on their quality by the technical committee and reviewers. Papers that describe on-going research and experimentation are encouraged. All paper submissions will be handled electronically and detailed instructions on submission procedure are available on IJCSI website (www.IJCSI.org).

For more information, please visit the journal website (www.IJCSI.org)

© IJCSI PUBLICATION 2013

www.IJCSI.org



The International Journal of Computer Science Issues (IJCSI) is a well-established and notable venue for publishing high quality research papers as recognized by various universities and international professional bodies. IJCSI is a refereed open access international journal for publishing scientific papers in all areas of computer science research. The purpose of establishing IJCSI is to provide assistance in the development of science, fast operative publication and storage of materials and results of scientific researches and representation of the scientific conception of the society.

It also provides a venue for researchers, students and professionals to submit ongoing research and developments in these areas. Authors are encouraged to contribute to the journal by submitting articles that illustrate new research results, projects, surveying works and industrial experiences that describe significant advances in field of computer science.

Indexing of IJCSI

1. Google Scholar
2. Bielefeld Academic Search Engine (BASE)
3. CiteSeerX
4. SCIRUS
5. Docstoc
6. Scribd
7. Cornell's University Library
8. SciRate
9. ScientificCommons
10. DBLP
11. EBSCO
12. ProQuest